

FIGURE 1

GGCTGAGGGGAGGCCCGAGCCTTTCTGGGGCCTGGGGGATCCTCTTGCACTGGTGGGTGGA
GAGAAGCGCCTGCAGCCAACCAGGGTCAGGCTGTGCTCACAGTTTCCTCTGGCGGCATGTAA
AGGCTCCACAAAGGAGTTGGGAGTTCAAATGAGGCTGCTGCGGACGGCCTGAGGATGGACCC
CAAGCCCTGGACCTGCCGAGCGTGGCACTGAGGCAGCGGCTGACGCTACTGTGAGGGAAAGA
AGGTTGTGAGCAGCCCCGAGGACCCCTGGCCAGCCCTGGCCCCAGCCTCTGCCGGAGCCCT
CTGTGGAGGCAGAGCCAGTGGAGCCCACTGAGGCAGGGCTGCTTGGCAGCCACCGGCCTGCA
ACTCAGGAACCCCTCCAGAGGCCATGGACAGGCTGCCCCGCTGACGGCCAGGGTGAAGCATG
TGAGGAGCCGCCCCGGAGCCAAGCAGGAGGGAAGAGGCTTTCATAGATTCTATTACAAAGA
ATAACCACCATTTTTGCAAGGACC**ATG**AGGCCACTGTGCGTGACATGCTGGTGGCTCGGACTG
CTGGCTGCCATGGGAGCTGTTGCAAGGCCAGGAGGACGGTTTTGAGGGCACTGAGGAGGGCTC
GCCAAGAGAGTTTCAATTTACCTAAACAGGTACAAGCGGGCGGGCGAGTCCAGGACAAGTGCA
CCTACACCTTCATTGTGCCCCAGCAGCGGGTCACGGGTGCCATCTGCGTCAACTCCAAGGAG
CCTGAGGTGCTTCTGGAGAACCAGTGCAATAAGCAGGAGCTAGAGCTGCTCAACAATGAGCT
GCTCAAGCAGAAGCGGCAGATCGAGACGCTGCAGCAGCTGGTGGAGGTGGACGGCGGCATTG
TGAGCGAGGTGAAGCTGCTGCGCAAGGAGAGCCGCAACATGAACTCGCGGGTCACGCAGCTC
TACATGCAGCTCCTGCACGAGATCATCCGCAAGCGGGACAACGCGTTGGAGCTCTCCAGCT
GGAGAACAGGATCCTGAACCAGACAGCCGACATGCTGCAGCTGGCCAGCAAGTACAAGGACC
TGGAGCACAAGTACCAGCACCTGGCCACACTGGCCCACAACCAATCAGAGATCATCGCGCAG
CTTGAGGAGCACTGCCAGAGGGTGCCCTCGGCCAGGCCCGTCCCCCAGCCACCCCCCGCTGC
CCCGCCCCGGGTCTACCAACCACCCACCTACAACCGCATCATCAACCAGATCTCTACCAACG
AGATCCAGAGTGACCAGAACCTGAAGGTGCTGCCACCCCTCTGCCCACTATGCCCACTCTC
ACCAGCCTCCCATCTTCCACCGACAAGCCGTCGGGGCCATGGAGAGACTGCCTGCAGGCCCT
GGAGGATGGCCACGACACCAGCTCCATCTACCTGGTGAAGCCGGAGAACACCAACCGCCTCA
TGCAGGTGTGGTGCAGACAGAGACACGACCCCGGGGGCTGGACCGTCATCCAGAGACGCCTG
GATGGCTCTGTAACTTCTTCAGGAAGTGGGAGACGTACAAGCAAGGGTTTGGGAACATTGA
CGGCGAATACTGGCTGGGCCTGGAGAACATTTACTGGCTGACGAACCAAGGCAACTACAAAC
TCCTGGTGACCATGGAGGACTGGTCCGGCCGCAAAGTCTTTGCAGAATACGCCAGTTTCCGC
CTGGAACCTGAGAGCGAGTATTATAAGCTGCGGCTGGGGCGCTACCATGGCAATGCGGGTGA
CTCCTTTACATGGCACAACGGCAAGCAGTTACACCACCTGGACAGAGATCATGATGTCTACA
CAGGAAACTGTGCCCACTACCAGAAGGGAGGCTGGTGGTATAACGCCTGTGCCCACTCCAAC
CTCAACGGGGTCTGGTACCGCGGGGGCCATTACCGGAGCCGCTACCAGGACGGAGTCTACTG
GGCTGAGTTCCGAGGAGGCTCTTACTCACTCAAGAAAGTGGTGATGATGATCCGACCGAACC
CCAACACCTTCCAC**TAA**GCCAGCTCCCCCTCCTGACCTCTCGTGGCCATTGCCAGGAGCCCA
CCCTGGTCAAGCTGGCCACAGCACAAAGAACAACCTCCTCACCAGTTTCATCCTGAGGCTGGGA
GGACCGGGATGCTGGATTCTGTTTTCCGAAGTCACTGCAGCGGATGATGGAAGTGAATCGAT
ACGGTGTTTTCTGTCCCTCCTACTTTTCTTACACCAGACAGCCCTCATGTCTCCAGGACA
GGACAGGACTACAGACAACCTTTCTTTAAATAAATTAAGTCTCTACAATAAAAAAAA

FIGURE 2

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA22779
><subunit 1 of 1, 493 aa, 1 stop
><MW: 57104, pI: 7.67, NX(S/T): 2
MRPLCVTCWWLGLLAAMGAVAGQEDGFEGTEEGSPREFIYLNRYKRAGESQDKCTYTFIVPQ
QRVTGAICVNSKEPEVLLLENRVHKQELELLNNELLKQKRQIETLQQLVEVDGGIVSEVKLLR
KESRNMNSRVLTQLYMQLLHEIIRKRDNALELSQLENRIILNQTADMLQLASKYKDLEHKYQHL
ATLAHNQSEIIAQLEEHCQRVPSARVPVQPPPAAPPRVYQPPTYNRIINQISTNEIQSDQNL
KVLPPPLPTMPTLTSLPSSTDKPSGPWRDCLQALEDGHDTSSIYLVKPEENTNRLMQVWCDQR
HDPGGWTVIQRRLDGSVNFNRNWETYKQGFGNIDGEYWLGLENIYWLTNQGNKLLVTMEDW
SGRKVFAEYASFRLEPESEYYKLRLGRYHGNAGDSFTWHNGKQFTTLDRDHDVYTGNCAYHQ
KGGWWYNACAHSNLNGVWYRGGHYRSRYQDGVYWAEFRGGSYSLLKKVMMIRPNPNTFH
```

Important features of the protein:

Signal peptide:

amino acids 1-22

N-glycosylation sites.

amino acids 164-168, 192-196

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 124-128

Tyrosine kinase phosphorylation sites.

amino acids 177-184, 385-393, 385-394, 461-468

N-myristoylation sites.

amino acids 12-18, 18-24, 22-28, 29-35, 114-120, 341-347,
465-471, 473-479

Amidation site.

amino acids 373-377

Fibrinogen beta and gamma chains C-terminal domain signature.

amino acids 438-451

Fibrinogen beta and gamma chains C-terminal domain proteins.

amino acids 305-343, 365-402, 411-424, 428-458

Trehalase proteins.

amino acids 275-292

3/75

FIGURE 3

CCCACGCGTCCGGCGCCGTGGCCTCGCGTCCATCTTTGCCGTTCTCTCGGACCTGTCACAAA
GGAGTCGCGCCGCGCCGCGCCGCCCCCTCCCTCCGGTGGGCCCGGGAGGTAGAGAAAGTCAGT
GCCACAGCCCGACCGCGCTGCTCTGAGCCCTGGGCACGCGGAACGGGAGGGAGTCTGAGGGT
TGGGGACGTCTGTGAGGGAGGGGAACAGCCGCTCGAGCCTGGGGCGGGCGGACCGGACTGGG
GCCGGGGTAGGCTCTGGAAAGGGCCCGGGAGAGAGGTGGCGTTGGTCAGAACCTGAGAAACA
GCCGAGAGGTTTTCCACCGAGGCCCCGCGCTTGAGGGATCTGAAGAGGTTCCCTAGAAGAGGGT
GTTCCCTCTTTTCGGGGGTCTCACCAGAAGAGGTTCTTGGGGGTCGCCCTTCTGAGGAGGCT
GCGGCTAACAGGGCCCAGAACTGCCATTGGATGTCCAGAATCCCCTGTAGTTGATAATGTTG
GGAATAAGCTCTGCAACTTTCTTTGGCATTCAGTTGTTAAAAACAAATAGGATGCAAATTCC
TCAACTCCAGGTTATGAAAACAGTACTTGGAAAACGAAAACCTACCTAAATGATCGTCTTTG
GTTGGGCCGTGTTCTTAGCGAGCAGAAGCCTTGGCCAGGGTCTGTTGTTGACTCTCGAAGAG
CACATAGCCCACTTCCTAGGGACTGGAGGTGCCGCTACTACCATGGGTAATTCCTGTATCTG
CCGAGATGACAGTGGAACAGATGACAGTGTGACACCCAACAGCAACAGGCCGAGAACAGTG
CAGTACCCACTGCTGACACAAGGAGCCAACCACGGGACCCTGTTTCGGCCACCAAGGAGGGGC
CGAGGACCTCATGAGCCAAGGAGAAAGAAACAAAATGTGGATGGGCTAGTGTGACACACT
GGCAGTAATACGGACTCTTGTAGATAAGTAAGTATCTGACTCACGGTCACCTCCAGTGGAAT
GAAAAGTGTTCTGCCCCGGAACCATGACTTTAGGACTCCTTCAGTTCCTTTAGGACATACTCG
CCAAGCCTTGTGCTCACAGGGCAAAGGAGAATATTTTAATGCTCCGCTGATGGCAGAGTAAA
TGATAAGATTTGATGTTTTTGCTTGCTGTCATCTACTTTGTCTGGAAATGTCTAAATGTTTC
TGTAGCAGAAAACACGATAAAGCTATGATCTTTATTAGAG

amino acids 18-24, 32-38, 34-40, 35-41, 51-57

CCCACGCGTCCGCGCAGTCGCGCAGTTCTGCCTCCGCCTGCCAGTCTCGCCCCGCGATCCCCGG
CCCGGGGGCTGTGGCGTCGACTCCGACCCAGGCAGCCAGCAGCCCGCGCGGGAGCCGGACCCG
CGCCGGAGGAGCTCGGACGGCATGCTGAGCCCCCTCCTTTGCTGAAGCCCGAGTGCGGAGAA
GCCCCGGGCAAACGCAGGCTAAGGAGACCAAAGCGGCGAAGTCGCGAGACAGCGGACAAGCAG
CGGAGGAGAAGGAGGAGGAGGCGAACCCAGAGAGGGGCAGCAAAGAAGCGGTGGTGGTGGG
CGTCGTGGCC**ATG**GCGGCGGCTATCGCCAGCTCGCTCATCCGTCAGAAGAGGCAAGCCCGCG
AGCGCGAGAAATCCAACGCCTGCAAGTGTGTCAGCAGCCCCAGCAAAGGCAAGACCAGCTGC
GACAAAAACAAGTTAAATGTCTTTTCCCGGGTCAAACCTCTTCGGCTCCAAGAAGAGGCGCAG
AAGAAGACCAGAGCCTCAGCTTAAGGGTATAGTTACCAAGCTATACAGCCGACAAGGCTACC
ACTTGCAGCTGCAGGCGGATGGAACCATTGATGGCACCAAAGATGAGGACAGCACTTACACT
CTGTTTAACCTCATCCCTGTGGGTCTGCGAGTGGTGGCTATCCAAGGAGTTCAAACCAAGCT
GTACTTGGCAATGAACAGTGAGGGATACTTGTACACCTCGGAACTTTTACACCTGAGTGCA
AATTCAAAGAATCAGTGTTTGAAAATTATTATGTGACATATTCATCAATGATATACCGTCAG
CAGCAGTCAGGCCGAGGGTGGTATCTGGGTCTGAACAAAGAAGGAGAGATCATGAAAGGCAA
CCATGTGAAGAAGAACAAGCCTGCAGCTCATTTTCTGCCTAAACCACTGAAAGTGGCCATGT
ACAAGGAGCCATCACTGCACGATCTCACGGAGTTCTCCCGATCTGGAAGCGGGACCCCAACC
AAGAGCAGAAGTGTCTCTGGCGTGCTGAACGGAGGCAAATCCATGAGCCACAATGAATCAAC
GTAGCCAGTGAGGGGCAAAGAAGGGCTCTGTAACAGAACCTTACCTCCAGGTGCTGTTGAAT
TCTTCTAGCAGTCCTTCACCCAAAAGTTCAAATTTGTGAGTGACATTTACCAAACAAACAGG
CAGAGTTCATATTCTATCTGCCATTAGACCTTCTTATCATCCATACTAAAGC

6/75

FIGURE 6

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA28498
><subunit 1 of 1, 245 aa, 1 stop
><MW: 27564, pI: 10.18, NX(S/T): 1
MAAAIASSLIRQKRQAREREKSNACKCVSSPSKGTSCDKNKLNVSFVKLFGSKKRRRRRP
EPQLKGIVTKLYSRQGYHLQLQADGTIDGTDKEDSTYTLFNLI PVGLRVVAIQGVQTKLYLA
MNSEG YLYTSELFTPECKFKESVFENYYVTYSSMIYRQQQSGRGWYLGLNKEGEIMKGNHVK
KNKPAAHFLPKPLKVAMYKEPSLHDLTEFSRSGSGTPTKSRSVSGVLNGGKSMHNEST
```

N-glycosylation site.

amino acids 242-246

Glycosaminoglycan attachment site.

amino acids 165-169, 218-222

Tyrosine kinase phosphorylation site.

amino acids 93-100

N-myristoylation site.

amino acids 87-93, 231-237

ATP/GTP-binding site motif A (P-loop).

amino acids 231-239

HBGF/FGF family proteins

amino acids 78-94, 102-153

FIGURE 7

ATGGCCGCGGCCATCGCTAGCGGCTTGATCCGCCAGAAGCGGCAGGCGCGGGAGCAGCACTG
GGACCGGCCGTCTGCCAGCAGGAGGCGGAGCAGCCCCAGCAAGAACCGCGGGCTCTGCAACG
GCAACCTGGTGGATATCTTCTCCAAAGTGCGCATCTTCGGCCTCAAGAAGCGCAGGTTGCGG
CGCCAAGATCCCCAGCTCAAGGGTATAGTGACCAGGTTATATTGCAGGCAAGGCTACTACTT
GCAAATGCACCCCGATGGAGCTCTCGATGGAACCAAGGATGACAGCACTAATTCTACACTCT
TCAACCTCATACCAGTGGGACTACGTGTTGTTGCCATCCAGGGAGTGAAAACAGGGTTGTAT
ATAGCCATGAATGGAGAAGGTTACCTCTACCCATCAGAACTTTTTACCCCTGAATGCAAGTT
TAAAGAATCTGTTTTTTGAAAATTATTATGTAATCTACTCATCCATGTTGTACAGACAACAG
AATCTGGTAGAGCCTGGTTTTTTGGGATTAAATAAGGAAGGGCAAGCTATGAAAGGGAACAGA
GTAAAGAAAACCAAACCAGCAGCTCATTTTCTACCCAAGCCATTGGAAGTTGCCATGTACCG
AGAACCATCTTTGCATGATGTTGGGGAAACGGTCCCGAAGCCTGGGGTGACGCCAAGTAAAA
GCACAAGTGCGTCTGCAATAATGAATGGAGGCAAACCAGTCAACAAGAGTAAGACAACA**TAG**

8/75

FIGURE 8

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA28503
><subunit 1 of 1, 247 aa, 1 stop
><MW: 27702, pI: 10.36, NX(S/T): 2
MAAAIASGLIRQKRQAREQHWDPRPSASRRRSSPSKNRGLCNGNLVDIFSKVRI FGLKKRRLR
RQDPQLKGIVTRLYCRQGYLQMHDPDGDGTDKDDSTNSTLFNLI PVGLRVVAIQGVKTGLY
IAMNGEGYLYPSELF TPECKFKESVFENYYVIYSSMLYRQQESGRAWFLGLNKEGQAMKGNR
VKKTKPAAHFLPKPLEVAMYREPSLHDVGETV PKPGVTPSKSTSASAIMNGGKPVNKS KTT
```

N-glycosylation site.

amino acids 100-104, 242-246

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 28-32, 29-33

Tyrosine kinase phosphorylation site.

amino acids 199-207

N-myristoylation site.

amino acids 38-44, 89-95, 118-124, 122-128, 222-228

HBGF/FGF family proteins.

amino acids 104-155, 171-198

FIGURE 9

CTCGCAGCCGAGCGCGGCCGGGGAAGGGCTCTCCTTCCAGCGCCGAGCACTGGGCCCCTGGCA
GACGCCCCAAGATTGTTGTGAGGAGTCTAGCCAGTTGGTGAGCGCTGTAATCTGAACCAGCT
GTGTCCAGACTGAGGCCCCATTTGCATTGTTTAACATACTTAGAAAATGAAGTGTTTCATTTT
TAACATTCTCTCTCCAATTGGTTTAAATGCTGAATTACTGAAGAGGGCTAAGCAAAACCAGGT
GCTTGCGCTGAGGGCTCTGCAGTGGCTGGGAGGACCCCGGCGCTCTCCCCGTGTCCTCTCCA
CGACTCGCTCGGCCCCCTCTGGAATAAAACACCCGCGAGCCCCGAGGGCCCAGAGGAGGCCGA
CGTGCCCGAGCTCCTCCGGGGGTCCCGCCCCGCGAGCTTTCTTCTCGCCTTCGCATCTCCTCC
TCGCGCGTCTTGGACATGCCAGGAATAAAAAGGATACTCACTGTTACCATTCTGGCTCTCTG
TCTTCCAAGCCCTGGGAATGCACAGGCACAGTGCACGAATGGCTTTGACCTGGATCGCCAGT
CAGGACAGTGTTTAGATATTGATGAATGCCGAACCATCCCCGAGGCCTGCCGAGGAGACATG
ATGTGTGTTAACCAAAATGGCGGGTATTTATGCATTCCCCGGACAAACCTGTGTATCGAGG
GCCCTACTCGAACCCTACTCGACCCCTACTCAGGTCCGTACCCAGCAGCTGCCCCACCAC
TCTCAGCTCCAACTATCCACGATCTCCAGGCCTCTTATATGCCGCTTTGGATACCAGATG
GATGAAAGCAACCAATGTGTGGATGTGGACGAGTGTGCAACAGATTCCCACCAGTGCAACCC
CACCCAGATCTGCATCAATACTGAAGGCGGGTACACCTGCTCCTGCACCGACGGATATTGGC
TTCTGGAAGGCCAGTGCTTAGACATTGATGAATGTGCGCTATGGTTACTGCCAGCAGCTCTGT
GCGAATGTTCTTGATCCTATTCTTGTACATGCAACCCTGGTTTTACCCTCAATGAGGATGG
AAGGTCTTGCCAAGATGTGAACGAGTGTGCCACCGAGAACCCTGCGTGCAAACCTGCGTCA
ACACCTACGGCTCTCTCATCTGCCGCTGTGACCCAGGATATGAACTTGAGGAAGATGGCGTT
CATTGCAGTGATATGGACGAGTGCAGCTTCTCTGAGTTCCTCTGCCAACATGAGTGTGTGAA
CCAGCCCGGCACATACTTCTGCTCCTGCCCTCCAGGCTACATCCTGCTGGATGACAACCGAA
GCTGCCAAGACATCAACGAATGTGAGCACAGGAACCACACGTGCAACCTGCAGCAGACGTGC
TACAATTTACAAGGGGGCTTCAAATGCATCGACCCCATCCGCTGTGAGGAGCCTTATCTGAG
GATCAGTGATAACCGCTGTATGTGTCTGCTGAGAACCCTGGCTGCAGAGACCAGCCCTTTA
CCATCTTGTACCGGGACATGGACGTGGTGTGAGGACGCTCCGTTCCCGCTGACATCTTCCAA
ATGCAAGCCACGACCCGCTACCCTGGGGCCTATTACATTTTCCAGATCAAATCTGGGAATGA
GGGAGAGAAATTTTACATGCGGCAAACGGGGCCCCATCAGTGCCACCCTGGTGATGACACGCC
CCATCAAAGGGCCCCGGGAAATCCAGCTGGACTTGGAATGATCACTGTCAACACTGTCATC
AACTTCAGAGGCAGCTCCGTGATCCGACTGCGGATATATGTGTGCGAGTACCCATTCTGAGC
CTCGGGCTGGAGCCTCCGACGCTGCCTCTCATTGGCACCAAGGGACAGGAGAAGAGAGGAAA
TAACAGAGAGAATGAGAGCGACACAGACGTTAGGCATTTCTGCTGAACGTTTCCCCGAAGA
GTCAGCCCCGACTTCCTGACTCTCACCTGTACTATTGCAGACCTGTACCCCTGCAGGACTTG
CCACCCCCAGTTCCCTATGACACAGTTATCAAAAAGTATTATCATTGCTCCCCTGATAGAAGA
TTGTTGGTGAATTTTCAAGGCCTTCAGTTTATTTCCACTATTTTCAAAGAAAATAGATTAGG
TTTGCGGGGGTCTGAGTCTATGTTCAAAGACTGTGAACAGCTTGCTGTCACTTCTTCACCTC
TTCCACTCCTTCTCTCACTGTGTTACTGCTTTGCAAAGACCCGGGAGCTGGCGGGGAACCT
GGGAGTAGCTAGTTTGCTTTTTTGCGTACACAGAGAAGGCTATGTAAACAAACCACAGCAGGA
TCGAAGGGTTTTTLAGAGAATGTGTTTCAAACCATGCCTGGTATTTTCAACCATAAAAGAAG
TTTCAGTTGTCCTTAAATTTGTATAACGGTTTAAATCTGTCTTGTTTCATTTTGAGTATTTT
AAAAAATATGTCGTAGAATTCCTTCGAAAGGCCTTCAGACACATGCTATGTTCTGTCTTCCC
AAACCCAGTCTCCTCTCCATTTTAGCCAGTGTTTTCTTTGAGGACCCCTTAATCTTGCTTT
CTTTAGAATTTTTACCCAATTGGATTGGAATGCAGAGGTCTCCAAACTGATTAAATATTTGA
AGAGA

10/75

FIGURE 10

MPGIKRILTVTILALCLPSPGNAQAQCTNGFDLDRQSGQCLDIDECRTIPEACRGDMMCVNQ
NGGYLCIPRTNPVYRGYPYSNPYSTPYSGPYPAAPPLSAPNYPTISRPLICRFGYQMDESNO
CVDVDECATDSHQCNPTQICINTEGGYTCSCDGYWLLLEGQCLDIDECRYGYCQQLCANVPG
SYSCTCNPGFTLNEDGRSCQDVNECATENPCVQTCVNTYGS LICRCDPGYELEEDGVHCSDM
DECSFSEFLCQHECVNQPGTYFCSCPPGYILLDDNRSCQDINECEHRNHTCNLQQTCYNLQG
GFKCIDPIRCEEPYLRISDNRCMCPAENPGCRDQPFITILYRDMDVVSGRSVPADIFQMQATT
RYPGAYYIFQIKSGNEGREFYMRQTGPISATLVMTRPIKGPREIQLDLEMITVNTVINFRGS
SVIRLRIYVSQYPF

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 283-287, 296-300

N-myristoylation sites.

amino acids 21-27, 64-70, 149-155, 186-192, 226-232, 242-248,
267-273, 310-316

Aspartic acid and asparagine hydroxylation sites.

amino acids 144-156, 181-193, 262-274

Cell attachment sequence.

amino acids 54-57

Calcium-binding EGF-like.

amino acids 131-166, 172-205, 211-245, 251-286

FIGURE 11

CAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCTCGGGGATCCTCTAGAGATCCCTC
GACCTCGACCCACGCGTCCGAACACAGGTCCTTGTGTGCTGCAGAGAAGCAGTTGTTTTGCTG
GAAGGAGGGAGTGCAGCGGGCTGCCCCGGGCTCCTCCCTGCCGCCTCCTCTCAGTGGATGGTT
CCAGGCACCCTGTCTGGGGCAGGGAGGGCACAGGCCTGCACATCGAAGGTGGGGTGGGACCA
GGCTGCCCCCTCGCCCCAGCATCCAAGTCCTCCCTTGGGCGCCCGTGGCCCTGCAGACTCTCA
GGGCTAAGGTCCTCTGTGCTTTTTGTTTCCACCTTAGAAGAGGCTCCGCTTGACTAAGAGT
AGCTTGAAGGAGGCACCA**ATG**CAGGAGCTGCATCTGCTCTGGTGGGCGCTTCTCCTGGGCCTG
GCTCAGGCCTGCCCTGAGCCCTGCGACTGTGGGGAAAAGTATGGCTTCCAGATCGCCGACTG
TGCCTACCGCGACCTAGAATCCGTGCCGCCTGGCTTCCCGGCCAATGTGACTACACTGAGCC
TGTCAGCGCAACCGGCTGCCAGGCTTGCCGGAGGGTGCCTTCAGGGAGGTGCCCTGCTGCAG
TCGCTGTGGCTGGCACACAATGAGATCCGCACGGTGGCCGCCGAGCCCTGGCCTCTCTGAG
CCATCTCAAGAGCCTGGACCTCAGCCACAATCTCATCTCTGACTTTGCCTGGAGCGACCTGC
ACAACCTCAGTGCCCTCCAATTGCTCAAGATGGACAGCAACGAGCTGACCTTCATCCCCCGC
GACGCCTTCCGCAGCCTCCGTGCTCTGCGCTCGCTGCAACTCAACCACAACCGCTTGCACAC
ATTGGCCGAGGGCACCTTCACCCCGCTCACCGCGCTGTCCCACCTGCAGATCAACGAGAACC
CCTTCGACTGCACCTGCGGCATCGTGTGGCTCAAGACATGGGCCCTGACCACGGCCGTGTCC
ATCCCGGAGCAGGACAACATCGCCTGCACCTCACCCCATGTGCTCAAGGGTACACCGCTGAG
CCGCCTGCCGCCACTGCCATGCTCGGCGCCCTCAGTGCAGCTCAGCTACCAACCCAGCCAGG
ATGGTGCCGAGCTGCGGCCTGGTTTTGTGCTGGCACTGCACTGTGATGTGGACGGGCAGCCG
GCCCCCTCAGCTTCACTGGCACATCCAGATAACCCAGTGGCATTGTGGAGATCACCAGCCCCAA
CGTGGGCACTGATGGGCGTGCCCTGCCTGGCACCCCTGTGGCCAGCTCCCAGCCGCGCTTCC
AGGCCTTTGCCAATGGCAGCCTGCTTATCCCCGACTTTGGCAAGCTGGAGGAAGGCACCTAC
AGCTGCCTGGCCACCAATGAGCTGGGCAGTGCTGAGAGCTCAGTGGACGTGGCACTGGCCAC
GCCCCGTGAGGGTGGTGAGGACACACTGGGGCGCAGGTTCCATGGCAAAGCGGTTGAGGGAA
AGGGCTGCTATACGGTTGACAACGAGGTGCAGCCATCAGGGCCGGAGGACAATGTGGTCATC
ATCTACCTCAGCCGTGCTGGGAACCCCTGAGGCTGCAGTCGCAGAAGGGGTCCCTGGGCAGCT
GCCCCCAGGCCTGCTCCTGCTGGGCCAAAGCCTCCTCCTCTTCTTCTTCTCCTCCTCCTCT**T**
AGCCCCACCCAGGGCTTCCCTAACTCCTCCCTTGCCCCCTACCAATGCCCCCTTAAGTGCTG
CAGGGGTCTGGGGTTGGCAACTCCTGAGGCCTGCATGGGTGACTTCACATTTTCTACCTCT
CCTTCTAATCTCTTCTAGAGCACCTGCTATCCCCAACTTCTAGACCTGCTCCAAACTAGTGA
CTAGGATAGAATTTGATCCCCCTAACTCACTGTCTGCGGTGCTCATTGCTGCTAACAGCATTG
CCTGTGCTCTCCTCTCAGGGGCAGCATGCTAACGGGGCGACGTCCTAATCCAACCTGGGAGAA
GCCTCAGTGGTGGAATTCAGGCACTGTGACTGTCAAGCTGGCAAGGGCCAGGATTGGGGGA
ATGGAGCTGGGGCTTAGCTGGGAGGTGGTCTGAAGCAGACAGGGAATGGGAGAGGAGGATGG
GAAGTAGACAGTGGCTGGTATGGCTCTGAGGCTCCCTGGGGCCTGCTCAAGCTCCTCCTGCT
CCTTGCTGTTTTCTGATGATTTGGGGGCTTGGGAGTCCCTTTGTCTCATCTGAGACTGAAA
TGTGGGGATCCAGGATGGCCTTCCTTCCTCTTACCCTTCCTCCCTCAGCCTGCAACCTCTAT
CCTGGAACCTGTCTCCCTTTCTCCCCAACTATGCATCTGTTGTCTGCTCCTCTGCAAAGGC
CAGCCAGCTTGGGAGCAGCAGAGAAATAAACAGCATTTCTGATGCCAAAAAAAAAAAAAAAAA
AAGGGCGGCCGCGACTCTAGAGTCGACCT

12/75

FIGURE 12

MQELHLLWWALLLGLAQACPEPCDCGEKYGFQIADCAYRDLESVPPGFPANVTTLSSLNRL
PGLPEGAFREVPLLQSLWLAHNEIRTVAAGALASLSHLKSLDLSHNLISDFAWSDLHNLSAL
QLLKMSNELTFIPRDAFRSLRALRSLQLNHNRLHTLAEGTFTPLTALSHLQINENPFDCCTC
GIVWLKTWALTAVSIPQDNIACTSPHVLKGTPLSRLPPLPCSAPSVQLSYQPSQDGAELR
PGFVLALHCDVDGQPAPQLHWHIQIPSGIVEITSPNVGTDGRALPGTPVASSQPRFQAFANG
SLLIPDFGKLEEGTYSCLATNELGSAESSVDVALATPGEGGEDTLGRRFHGKAVEGKGCYTV
DNEVQPSGPEDNVVVIYLSRAGNPEAAVAEGVPGQLPPGLLLLGQSLLLFFFLTSF

Important features of the protein:

Signal peptide:

amino acids 1-18

Transmembrane domain:

amino acids 403-418

N-glycosylation sites.

amino acids 51-55, 120-124, 309-313

Tyrosine kinase phosphorylation site.

amino acids 319-326

N-myristoylation sites.

amino acids 14-20, 64-70, 92-98, 218-224, 294-300, 323-329,
334-340, 350-356, 394-400

Amidation site.

amino acids 355-359

Leucine rich repeats.

amino acids 51-74, 75-98, 99-122, 123-146, 147-170

Leucine rich repeat C-terminal domain.

amino acids 180-230

CCAGGCCGGGAGGACGCGCCAGCCAGCGTCTAAACGGGAACAGCCCTGGCTGAGGGAGCTGC
AGCGCAGCAGAGTATCTGACGGCGCCAGGTGCGTAGGTGCGGCACGAGGAGTTTTCCCGGC
AGCGAGGAGGTCTTGAGCAGCATGGGCCGGAGGAGCGCCTTCCCTGCCGCCGCGCTCTGGCT
CTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGAGGCCGGGCCGCCGAGGAGGAGA
GCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCATAGGATTTGAAGAAGATATC
CTGATTGTTTCAGAGGGGAAAATGGCACCTTTTACACATGATTTCAGAAAAGCGCAACAGAG
AATGCCAGCTATTCTGTCAATATCCATTCCATGAATTTTACCTGGCAAGCTGCAGGGCAGG
CAGAATACTTCTATGAATTCTGTCTTGCCTGCGCTCCCTGGATAAAGGCATCATGGCAGATCCA
ACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCATCAGTTGTTCAAGTTGGTTT
CCCATGTCTTGGAACACAGGATGGGGTGGCAGCATTTGAAGTGGATGTGATTGTTATGAATT
CTGAAGGCAACACCATTCTCAAACACCTCAAATGCTATCTTCTTTAAACATGTCAACAA
GCTGAGTGCCCGAGGCGGGTGCCGAAATGGAGGCTTTTGTAATGAAAGACGCATCTGCGAGTG
TCCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTGTACCCACGATGTATGAATG
GTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCACCTGGATTCTATGGAGTGAAGTGT
GACAAAGCAAACCTGCTCAACCACCTGCTTTAATGGAGGGACCTGTTTCTACCCTGGAAAATG
TATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAGCAAATGCCACAACCCTGTC
GAAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTTCAAAGGTTACCAGGGAGAC
CTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACATGGAACCTGCCATGAACCCAA
CAAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAATAAAAGGTACGAAGCCAGCC
TCATACATGCCCTGAGGCCAGCAGGCGCCAGCTCAGGCAGCACACGCCTTCACTTAAAAAG
GCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGGTGAACTCCGACATCTGAAAC
GTTTTAAGTTACACCAAGTTCATAGCCTTTGTTAACCTTTTCATGTGTTGAATGTTCAAATAA
TGTTCAATTACACTTAAGAATACTGGCCTGAATTTTATTAGCTTCATTATAAATCACTGAGCT
GATATTTACTCTTCCTTTTAAGTTTCTAAGTACGTCTGTAGCATGATGGTATAGATTTTCT
TGTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGATCAGGTAAAATTTTCAGTG
TGTAGTTGGCAGATATTTTCAAATTAACAATGCATTTATGGTGTCTGGGGGCAGGGGAACAT
CAGAAAGGTTAAATTGGGCAAAAATGCGTAAGTCACAAGAATTTGGATGGTGCAGTTAATGT
TGAAGTTACAGCATTTTCAAGTTTTATTGTCAGATATTTAGATGTTTGTTACATTTTTTAAAAA
TTGCTCTTAATTTTTTAACTCTCAATACAATATATTTTGACCTTACCATTATTCCAGAGATT
CAGTATTAATAAAAAAAAAAATTACACTGTGGTAGTGGCATTTAACAATATAATATATTCTA
AACACAATGAAATAGGGAATATAATGTATGAACTTTTTGCATTGGCTTGAAGCAATATAATA
TATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTATACTGTTTGTATGTATAAAAT
AAAGGTGCTGCTTTTAGTTTTTTTGGAAAAA

FIGURE 14

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESLYLWIDAHQARVLIGFEEDILIVSEGK
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFSLRSLDKGIMADPTVNVPLL
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPGGC
RNGGFCNERRICECPDGFHGPHEKALCTPRCMNGGLCVTPGFCICPPPGFYGVNCDKANCST
TCFNGGTCFYPGKCICPPGLEGEQCEISKCPQPCRNGGKCIGKSKCKCSKGYQGDLCSKPVC
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAERRDP
PESNYIW

Signal sequence.

amino acids 1-28

N-glycosylation sites.

amino acids 88-92, 245-249

Tyrosine kinase phosphorylation site.

amino acids 370-378

N-myristoylation sites.

amino acids 184-190, 185-191, 189-195, 315-321

ATP/GTP-binding site motif A (P-loop).

amino acids 285-293

EGF-like domain cysteine pattern signature.

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

FIGURE 15

AGAACCTCAGAAATGTGAGTTATTTGGGAATGGCTGTTTGTAATGTCCTTACGTAAGCCAA
GAGGAGGTCTTGACTTGGGGTCCCAGGGGTACCGCAGATCCCAGGGACTGGAGCAGCACTAG
CAAGCTCTGGAGGATGAGCCAGGAGTCTGGAATTGAGGCTGAGCCAAAGACCCCAGGGCCGT
CTCAGTCTCATAAAAGGGGATCAGGCAGGAGGAGTTTGGGAGAAACCTGAGAAGGGCCTGAT
TTGCAGCATC**ATG**ATGGGCCTCTCCTTGGCCTCTGCTGTGCTCCTGGCCTCCCTCCTGAGTC
TCCACCTTGGAAGTGGCACACGTGGGAGTGACATATCCAAGACCTGCTGCTTCCAATACAGC
CACAAGCCCCTTCCCTGGACCTGGGTGCGAAGCTATGAATTCACCAGTAACAGCTGCTCCCA
GCGGGCTGTGATATTCACTACCAAAGAGGCAAGAAAGTCTGTACCCATCCAAGGAAAAAAT
GGGTGCAAAAATACATTTCTTTACTGAAAACCTCCGAAACAATTG**TGA**CTCAGCTGAATTTTC
ATCCGAGGACGCTTGGACCCCGCTCTTGGCTCTGCAGCCCTCTGGGGAGCCTGCGGAATCTT
TTCTGAAGGCTACATGGACCCGCTGGGGAGGAGAGGGTGTTTCCTCCCAGAGTTACTTTAAT
AAAGGTTGTTCATAGAGTTGAAA
AAA

16/75

FIGURE 16

MMGLSLASAVLLASLLSLHLGTATRGSDISKTC CFQYSHKPLPWTWVRSYEFTSNSCSQRAV
IFTTKRGKKVCTHPRKKWVQKYISLLKTPKQL

Important features of the protein:

Signal peptide:

amino acids 1-23

N-myristoylation sites.

amino acids 3-9, 26-32

Amidation site.

amino acids 68-72

Small cytokines (intecrine/chemokine).

amino acids 23-88

FIGURE 17

GCGAGAACCCTTTGCACGCGCACAAACTACGGGGACGATTTCTGATTGATTTTGGCGCTTTCGATCCACCCTCC
TCCCTTCTCATGGGGACTTTGGGGACAAAGCGTCCCGACCGCTCGAGCGCTCGAGCAGGGCGCTATCCAGGAGC
CAGGACAGCGTCGGGAACCAGACCATGGCTCCTGGACCCCAAGATCCTTAAGTTTCGTCGTCCTCATCGTCGCGG
TTCTGCTGCCGGTCCGGGTTGACTCTGCCACCATCCCCGGCAGGACGAAGTTCCCCAGCAGACAGTGGCCCCA
CAGCAACAGAGGCGCAGCCTCAAGGAGGAGGAGTGTCCAGCAGGATCTCATAGATCAGAATATACTGGAGCCTG
TAACCCGTGCACAGAGGGTGTGGATTACACCATTGTCTCCAACAATTTGCCTTCTTGCCTGCTATGTACAGTTT
GTAAATCAGGTCAAACAAATAAAAGTTCTGTACCACGACCAGAGACACCGTGTGTGTCAGTGTGAAAAAGGAAGC
TTCCAGGATAAAAACTCCCCTGAGATGTGCCGGACGTGTAGAACAGGGTGTCCAGAGGGATGGTCAAGGTCAG
TAATTGTACGCCCCGGAGTGACATCAAGTGCAAAAATGAATCAGCTGCCAGTTCCACTGGGAAAACCCAGCAG
CGGAGGAGACAGTGACCACCATCCTGGGGATGCTTGCCTCTCCCTATCACTACCTTATCATCATAGTGGTTTTA
GTCATCATTTTAGCTGTGGTTGTGGTTGGCTTTTCATGTCCGAAGAAATTCATTTCTTACCTCAAAGGCATCTG
CTCAGGTGGTGGAGGAGGTCCCGAACGTGTGCACAGAGTCTTTTCCGGCGGCGTTTCAGTCTCCTTACAGGTTT
CTGGGGCGGAGGACAATGCCCGCAACGAGACCCCTGAGTAACAGATACTTGCAGCCACCCAGGTCTCTGAGCAG
GAAATCCAAGGTCAGGAGCTGGCAGAGCTAACAGGTGTGACTGTAGAGTCCGAGAGGAGCCACAGCGTCTGCT
GGAACAGGCAGAAGCTGAAGGGTGTGAGGAGGAGGCTGCTGGTTCCAGTGAATGACGCTGACTCCGCTGACA
TCAGCACCTTGCTGGATGCCTCGGCAACACTGGAAGAAGGACATGCAAGGAACAATTCAGGACCAACTGGTG
GGCTCCGAAAAGCTCTTTTATGAAGAAGATGAGGCAGGCTCTGCTACGTCCTGCCTGTGAAAGAATCTCTTCAG
GAAACCAGAGCTTCCCTCATTTACCTTTTCTCTACAAAGGGAAGCAGCCTGGAAGAAACAGTCCAGTACTTGA
CCCATGCCCCAACAACTCTACTATCCAATATGGGGCAGCTTACCAATGGTCTCTAGAACTTTGTAAACGCACTT
GGAGTAATTTTTATGAAATACTGCGTGTGATAAGCAAACGGGAGAAATTTATATCAGATTCTTGGCTGCATAGT
TATACGATTGTGTATTAAGGGTCGTTTTAGGCCACATGCGGTGGCTCATGCCTGTAATCCAGCAGCTTTGATAG
GCTGAGGCAGGTGGATTGCTTGAGCTCGGGAGTTTGAAGACAGCCTCATCAACACAGTGAAGTCCATCTCAAT
TTAAAAAGAAAAAAGTGTTTTAGGATGTGATTCCTTTCAGTTCTTCATCATGAGACAAGTCTTTTTTCTGTC
TTCTTATATTGCAAGCTCCATCTCTACTGGTGTGTGCATTTAATGACATCTAACTACAGATGCCGCACAGCCAC
AATGCTTTGCCTTATAGTTTTTTAACTTTAGAACGGGATTATCTTGTATTACCTGTATTTTCAGTTTCGGATA
TTTTTGACTTAATGATGAGATTATCAAGACGTAGCCCTATGCTAAGTCATGAGCATATGGAATTACGAGGGTTT
GACTTAGAGTTTTGAGCTTTAAGATAGGATTATTTGGGGCTTACCCCCACCTTAATTAGAGAAACATTTATATTG
CTTACTACTGTAGGCTGTACATCTCTTTCCGATTTTGTATAATGATGTAAACATGGAAAACTTTAGGAAAT
GCACTTATTAGGCTGTTTACATGGGTTGCCTGGATACAAATCAGCAGTCAAAAATGACTAAAAATATAACTAGT
GACGGAGGGAGAAATCCTCCCTCTGTGGGAGGCACTTACTGCATTCCAGTTCTCCCTCCTGCGCCCTGAGACTG
GACCAGGGTTTGATGGCTGGCAGCTTCTCAAGGGGCAGCTTGTCTTACTTGTTAATTTTAGAGGTATATAGCCA
TATTTATTTATAAATAAATATTTATTTATTTATTTATAAGTAGATGTTTACATATGCCCAGGATTTTGAAGAGC
CTGGTATCTTTGGGAAGCCATGTGTCTGGTTTGTCTGCTGGGACAGTCATGGGACTGCATCTTCCGACTTGTG
CACAGCAGATGAGGACAGTGAGAATTAAGTTAGATCCGAGACTGCGAAGAGCTTCTCTTTCAAGCGCCATTACA
GTTGAACGTTAGTGAATCTTGAGCCTCATTTGGGCTCAGGGCAGAGCAGGTGTTTATCTGCCCCGGCATCTGCC
ATGGCATCAAGAGGGAAGAGTGGACGGTGTCTGGGAATGGTGTGAAATGGTTGCCGACTCAGGCATGGATGGGC
CCCTCTCGCTTCTGGTGGTCTGTGAAGTGAAGTCCCTGGGATGCCTTTTAGGGCAGAGATTCCTGAGCTGCGTTT
TAGGGTACAGATTCCCTGTTTGAAGGAGCTTGGCCCCCTCTGTAAGCATCTGACTCATCTCAGAGATATCAATTCT
TAAACACTGTGACAACGGGATCTAAATGGCTGACACATTTGTCTTGTGTACAGTTCATTTATTTTATTTAAA
AACCTCAGTAATCGTTTTAGCTTCTTTCCAGCAAACCTCTTCTCCACAGTAGCCAGTCGTGGTAGGATAAATTA
CGGATATAGTCATTCTAGGGGTTTCAGTCTTTTCCATCTCAAGGCATTGTGTGTTTTGTTCCGGGACTGGTTTG
GCTGGGACAAAGTTAGAAGTGCCTGAAGTTCGCACATTAGATTGTTGTGTCCATGGAGTTTTAGGAGGGGATG
GCCTTTCCGGTCTTCGCACTTCCATCCTCTCCCACTCCATCTGGCGTCCCACACCTTGTCCCCTGCACTTCTG
GATGACACAGGGTGTGCTGCCTCCTAGTCTTTGCCCTTGTGCGGCTTCTGTGCAGGAGACTTGGTCTCAAAG
CTCAGAGAGAGCCAGTCCGGTCCCAGCTCCTTTGTCCCTTCTCAGAGGCCTTCCCTGAAGATGCATCTAGACT
ACCAGCCTTATCAGTGTTTAAGCTTATTCCTTTAACATAAGCTTCCCTGACAACATGAAATGTTGGGGTTTTTT
GGCGTTGGTTGATTTGTTTAGGTTTTGCTTTATACCCGGGCCAAATAGCACATAACACCTGGTTATATATGAAA
TACTCATATGTTTATGACCAAAATAAATATGAAACCTCATRTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

FIGURE 18

MGLWGQSVPTASSARAGRYPGARTASGTRPWLLDPKILKFVVFIVAVLLPVRVDSATIPRQD
 EVPQQTVAPOQQRRSLKEEECPAGSHRSEYTGACNPCTEGVDYTIASNNLPSCLLCTVCKSG
 QTNKSSCTTTTRDTVQCCEKGSFQDKNSPEMCRTCRTGCPGMVKVSNCTPRSDIKCKNESAA
 SSTGKTPAAEETVTTILGMLASPHYLIIVVLVIILAVVVVVGFSRKKFISYLLKGICSGGG
 GGPERVHRVLFRRRSCPSRVPGAEDNARNETLSNRYLQPTQVSEQEIQGQELAEALTGVTVES
 PEEPQRLLEQAEAEGCQRRRLLVPVNDADSADISTLLDASATLEEGHAKETIQDQLVGSEKL
 FYEDEAGSATSCL

Important features of the protein:

Transmembrane domains:

amino acids 35-52, 208-230

N-glycosylation sites.

amino acids 127-131, 182-186, 277-281

Glycosaminoglycan attachment site.

amino acids 245-249

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 260-264

N-myristoylation sites.

amino acids 21-27, 86-92, 102-108, 161-167, 242-248, 270-276,
 297-303, 380-386

ATP/GTP-binding site motif A (P-loop).

amino acids 185-193

TNFR/NGFR cysteine-rich region.

amino acids 99-139

FIGURE 19

GCGGCACCTGGAAGATGCGCCCATTTGGCTGGTGGCCTGCTCAAGGTGGTGTTCGTGGTCTTC
GCCTCCTTGTGTGCCTGGTATTCGGGGTACCTGCTCGCAGAGCTCATTCCAGATGCACCCCT
GTCCAGTGCTGCCTATAGCATCCGCAGCATCGGGGAGAGGCCTGTCCTCAAAGCTCCAGTCC
CCAAAAGGCAAAAATGTGACCACTGGACTCCCTGCCCATCTGACACCTATGCCTACAGGTTA
CTCAGCGGAGGTGGCAGAAGCAAGTACGCCAAAATCTGCTTTGAGGATAACCTACTTATGGG
AGAACAGCTGGGAAATGTTGCCAGAGGAATAAACATTGCCATTGTCAACTATGTAAGTGGGA
ATGTGACAGCAACACGATGTTTTGATATGTATGAAGGCGATAACTCTGGACCGATGACAAAG
TTTATTCAGAGTGCTGCTCCAAAATCCCTGCTCTTCATGGTGACCTATGACGACGGAAGCAC
AAGACTGAATAACGATGCCAAGAATGCCATAGAAGCACTTGGAAGTAAAGAAATCAGGAACA
TGAAATTCAGGTCTAGCTGGGTATTTATTGCAGCAAAAGGCTTGGAACCTCCCTTCCGAAATT
CAGAGAGAAAAGATCAACCACTCTGATGCTAAGAACAACAGATATTCTGGCTGGCCTGCAGA
GATCCAGATAGAAGGCTGCATACCCAAAGAACGAAGCTTGACACTGCAGGGTCCTGAGTAAAT
GTGTTCTGTATAAACAAATGCAGCTGGAATCGCTCAAGAATCTTATTTTTCTAAATCCAACA
GCCCATATTTGATGAGTATTTTGGGTTTGTGTAAACCAATGAACATTTGCTAGTTGTATCA
AATCTTGGTACGCAGTATTTTATACCAGTATTTTATGTAGTGAAGATGTCAATTAGCAGGA
AACTAAAATGAATGGAAATTCTTAAAAAAAAA

20/75

FIGURE 20

MRPLAGGLLKVVVFVVFASLCAWYSGYLLAELIPDAPLSSAAYSIRSIGERPVLKAPVPKRQK
CDHWTPCPSDTYAYRLLSGGGRSKYAKICFEDNLLMGEQLGNVARGINIAIVNYVTGNVTAT
RCFDMYEGDNSGPMTKFIQSAAPKSLLFMVTYDDGSTRLNNDKNAIEALGSKEIRNMKFRS
SWVFIAAKGLELPSEIQREKINHSDAKNNRYSGWPAEIQIEGCI PKERS

Important features of the protein:

Signal peptide:

amino acids 1-20

N-glycosylation sites.

amino acids 120-124, 208-212

Glycosaminoglycan attachment site.

amino acids 80-84

N-myristoylation sites.

amino acids 81-87, 108-114, 119-125

21/75

FIGURE 21

CCGGGGAGGGGAGGGCCCGTCCCGCCCCCTCCCCGTCTCTCCCCGCCCCCTCCCCGTCCCTCCC
GCCGAAGCTCCGTCCCGCCCCGCGGGCCGGCTCCGCCCTCACCTCCCGGCCGCGGGCTGCCCTC
TGCCCGGGTTGTCCAAG**ATG**GAGGGCGCTCCACCGGGGTCGCTCGCCCTCCGGCTCCTGCTG
TTCGTGGCGCTACCCGCCTCCGGCTGGCTGACGACGGGCGCCCCCGAGCCGCCGCCGCTGTC
CGGAGCCCCACAGGACGGCATCAGAATTAATGTAACCTACACTGAAAGATGATGGGGACATAT
CTAAACAGCAGGTTGTTCTTAACATAACCTATGAGAGTGGACAGGTGTATGTAAATGACTTA
CCTGTAAATAGTGGTGTAAACCCGAATAAGCTGTCAGACTTTGATAGTGAAGAATGAAAATCT
TGAAAATTTGGAGGAAAAAGAATATTTTGGAAATTGTCAGTGTAAAGGATTTTAGTTCATGAGT
GGCCTATGACATCTGGTTCCAGTTTGCAACTAATTGTCATTCAAGAAGAGGTAGTAGAGATT
GATGGAAAACAAGTTCAGCAAAGGATGTCAGTGAATTTGATATTTTAGTTAAGAACCGGGG
AGTACTCAGACATTCAAACCTATACCCTCCCTTTGGAAGAAAGCATGCTCTACTCTATTTCTC
GAGACAGTGACATTTTATTTACCCTTCCTAACCTCTCCAAAAAAGAAAGTGTAGTTCACTG
CAAACCACTAGCCAGTATCTTATCAGGAATGTGGAAACCACTGTAGATGAAGATGTTTTACC
TGGCAAGTTACCTGAAACTCCTCTCAGAGCAGAGCCGCCATCTTCATATAAGGTAATGTGTC
AGTGGATGGAAAAGTTTAGAAAAGATCTGTGTAGGTTCTGGAGCAACGTTTTCCAGTATTC
TTTCAGTTTTTTGAACATCATGGTGGTTGGAATTACAGGAGCAGCTGTGGTAATAACCATCTT
AAAGGTGTTTTTCCCAGTTTCTGAATACAAAGGAATTCTTCAGTTGGATAAAGTGGACGTCA
TACCTGTGACAGCTATCAACTTATATCCAGATGGTCCAGAGAAAAGAGCTGAAAACCTTGAA
GATAAAACATGTATT**TAA**AACGCCATCTCATATCATGGACTCCGAAGTAGCCTGTTGCCTCC
AAATTTGCCACTTGAATATAATTTTCTTTAAATCGTT

22/75

FIGURE 22

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60783
><subunit 1 of 1, 330 aa, 1 stop
><MW: 36840, pI: 4.84, NX(S/T): 4
MEGAPPGSLALRLLLFVALPASGWLTTGAPEPPPLSGAPQDGIRINVTTLKDDGDISKQQVV
LNITYESGQVYVNDLPVNSGVTRISCQTLIVKNENLENLEEKEYFGIVSVRILVHEWPMTSG
SSLQLIVIQEEVVEIDGKQVQKQDVTEIDILVKNRGVLRHSNYTLPLEESMLYSISRSDIL
FTLPNLSKKESVSSLQTTSQYLIRNVETTVDEDVLP GKLPETPLRAEPPSSYKVMCQWMEKF
RKDLCRFWSNVFPVFFQFLNIMVVGITGAADVITILKVFFPVSEYKGILQLDKVDVIPVTAI
NLYPDGPPEKRAENLEDKTCI
```

Important features of the protein:

Signal peptide:

amino acids 1-23

Transmembrane domain:

amino acids 266-284

Leucine zipper pattern.

amino acids 155-176

N-glycosylation sites.

amino acids 46-49, 64-67, 166-169, 191-194

23/75

FIGURE 23

CGTCTCTGCGTTTCGCC**ATG**CGTCCCGGGGCGCCAGGGCCACTCTGGCCTCTGCCCTGGGGGG
 CCCTGGCTTGGGCCGTGGGCTTCGTGAGCTCCATGGGCTCGGGGAACCCCGCGCCCGGTGGT
 GTTTGCTGGCTCCAGCAGGGCCAGGAGGCCACCTGCAGCCTGGTGCTCCAGACTGATGTCAC
 CCGGGCCGAGTGCTGTGCCTCCGGCAACATTGACACCGCCTGGTCCAACCTCACCCACCCGG
 GGAACAAGATCAACCTCCTCGGCTTCTTGGGCCTTGCTCCACTGCCTTCCCTGCAAAGATTCTG
 TCGGACGGCGTGGAGTGCGGCCCGGGCAAGGCGTGCCGCATGCTGGGGGGCCGCCCGCGCTG
 CGAGTGCGCGCCCGACTGCTCGGGGCTCCCGGCGCGGCTGCAGGTCTGCGGCTCAGACGGCG
 CCACCTACCGCGACGAGTGCGAGCTGCGCGCCGCGCTGCCGCGGCCACCCGGACCTGAGC
 GTCATGTACCGGGGCGCTGCCGCAAGTCCGTGTGAGCACGTGGTGTGCCCGCGGCCACAGTC
 GTGCGTCTGTGGACCAGACGGGCAGCGCCACTGCGTGGTGTGTGAGCGGCGCCCTGCCCTG
 TGCCCTCCAGCCCCGGCCAGGAGCTTTGCGGCAACAACAACGTCACCTACATCTCCTCGTGC
 CACATGCGCCAGGCCACCTGCTTCCCTGGGCGCTCCATCGGCGTGCGCCACGCGGGCAGCTG
 CGCAGGCACCCCTGAGGAGCCGCCAGGTGGTGAGTCTGCAGAAGAGGAAGAGAAGTTCGTG**T**
GAGCCTGCAGGACAGGCCTGGGCCTGGTGCCCGAGGCCCCCATCATCCCTGTTATTTATT
 GCCACAGCAGAGTCTAATTTATATGCCACGGACACTCCTTAGAGCCCGGATTCGGACCACTT
 GGGGATCCCAGAACCTCCCTGACGATATCCTGGAAGGACTGAGGAAGGGAGGCCTGGGGGCC
 GGCTGGTGGGTGGGATAGACCTGCGTTCCGGACACTGAGCGCCTGATTTAGGGCCCTTCTCT
 AGGATGCCCCAGCCCCTACCCTAAGACCTATTGCCGGGGAGGATTCCACACTTCCGCTCCTT
 TGGGGATAAACCTATTAATTATTGCTACTATCAAGAGGGCTGGGCATTCTCTGCTGGTAATT
 CCTGAAGAGGCATGACTGCTTTTCTCAGCCCCAAGCCTCTAGTCTGGGTGTGTACGGAGGGT
 CTAGCCTGGGTGTGTACGGAGGGTCTAGCCTGGGTGAGTACGGAGGGTCTAGCCTGGGTGAG
 TACGGAGGGTCTAGCCTGGGTGAGTACGGAGGGTCTAGCCTGGGTGTGTATGGAGGATCTAG
 CCTGGGTGAGTATGGAGGGTCTAGCCTGGGTGAGTATGGAGGGTCTAGCCTGGGTGTGTATG
 GAGGGTCTAGCCTGGGTGAGTATGGAGGGTCTAGCCTGGGTGTGTATGGAGGGTCTAGCCTG
 GGTGAGTATGGAGGGTCTAGCCTGGGTGTGTACGGAGGGTCTAGTCTGAGTGCCTGTGGGGA
 CCTCAGAACACTGTGACCTTAGCCCAGCAAGCCAGGCCCTTCATGAAGGCCAAGAAGGCTGC
 CACCATTCCCTGCCAGCCCAAGAACTCCAGCTTCCCCACTGCCTCTGTGTGCCCTTTGCGT
 CCTGTGAAGGCCATTGAGAAATGCCAGTGTGCCCCCTGGGAAAGGGCACGGCCTGTGCTCC
 TGACACGGGCTGTGCTTGGCCACAGAACCACCCAGCGTCTCCCCTGCTGCTGTCCACGTCAG
 TTCATGAGGCAACGTCGCGTGGTCTCAGACGTGGAGCAGCCAGCGGCAGCTCAGAGCAGGGC
 ACTGTGTCCGGCGGAGCCAAGTCCACTCTGGGGGAGCTCTGGCGGGGACCACGGGCCACTGC
 TCACCCACTGGCCCCGAGGGGGGTGTAGACGCCAAGACTCACGCATGTGTGACATCCGGAGT
 CCTGGAGCCGGGTGTCCAGTGGCACCACTAGGTGCCTGCTGCCTCCACAGTGGGGTTTACA
 CCCAGGGGCTCCTTGGTCCCCCAACAACCTGCCCCGGCCAGGCCTGCAGACCCAGACTCCAGCC
 AGACCTGCCTCACCCACCAATGCAGCCGGGGCTGGCGACACCAGCCAGGTGCTGGTCTTGGG
 CCAGTTCTCCACGACGGCTCACCCCTCCCTCCATCTGCGTTGATGCTCAGAATCGCCTACC
 TGTGCCTGCGTGTAAACCACAGCCTCAGACCAGCTATGGGGAGAGGACAACACGGAGGATAT
 CCAGCTTCCCCGGTCTGGGGTGAGGAATGTGGGGAGCTTGGGCATCCTCCTCCAGCCTCCTC
 CAGCCCCCAGGCAGTGCTTACCTGTGGTGCCAGAAAAGTGCCCCCTAGGTTGGTGGGTCTA
 CAGGAGCCTCAGCCAGGCAGCCACCCACCCCTGGGGCCCTGCCTCACCAAGGAAATAAAGA
 CTAAGCCATAAAAAAA

24/75

FIGURE 24

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62306

<subunit 1 of 1, 263 aa, 1 stop

<MW: 27663, pI: 6.77, NX(S/T): 2

MRPGAPGPLWPLPWGALAWAVGFVSSMSGGNPAPGGVCWLQQGQEATCSLVLQTDVTRA ECC
ASGNIDTAWSNLTHPGNKINLLGFLGLVHCLPCKDSCDGV ECGPGKACRMLGGRPRCECAPD
CSGLPARLQVCGSDGATYRDECELRAARCRGHPDLSVMYRGRCRKSC EHVVCPRPQSCVVDQ
TGAHCVVCRAAPCPVPSSPGQELCGNNNVTYISSCHMRQATCFLGRSIGVRHAGSCAGTPE
EPPGGESAEEEEENFV

Important features:

Signal peptide:

amino acids 1-20

N-glycosylation sites.

amino acids 73-77, 215-219

Osteonectin domain proteins.

amino acids 97-130, 169-202

25/75

FIGURE 25

TGCAGAGCTTGTGGAGGCC**ATG**GGGCGCGTCGTCGCGGAGCTCGTCTCCTCGCTGCTGGGGT
TGTGGCTGTTGCTGTGCAGCTGCGGATGCCCCGAGGGCGCCGAGCTGCGTGCTCCGCCAGAT
AAAATCGCGATTATTGGAGCCGGAATTGGTGGCACTTCAGCAGCCTATTACCTGCGGCAGAA
ATTTGGGAAAGATGTGAAGATAGACCTGTTTGAAAGAGAAGAGGTCGGGGGCCCGCTGGCTA
CCATGATGGTGCAGGGGCAAGAATACGAGGCAGGAGGTTCTGTTCATCCATCCTTTAAATCTG
CACATGAAACGTTTTGTCAAAGACCTGGGTCTCTCTGCTGTTTCAGGCCTCTGGTGGCCTACT
GGGGATATATAATGGAGAGACTCTGGTATTTGAGGAGAGCAACTGGTTCATAATTAACGTGA
TTAAATTAGTTTGGCGCTATGGATTTCAATCCCTCCGTATGCACATGTGGGTAGAGGACGTG
TTAGACAAGTTCATGAGGATCTACCGCTACCACTCTCATGACTATGCCTTCAGTAGTGTCGA
AAAATTACTTCATGCTCTAGGAGGAGATGACTTCCTTGGAATGCTTAATCGAACACTTCTTG
AAACCTTGCAAAAGGCCGGCTTTTCTGAGAAGTTCCTCAATGAAATGATTGCTCCTGTTATG
AGGGTCAATTATGGCCAAAGCACGGACATCAATGCCTTTGTGGGGGCGGTGTCACTGTCCTG
TTCTGATTCTGGCCTTTGGGCAGTAGAAGGTGGCAATAAACTTGTTTGCTCAGGGCTTCTGC
AGGCATCCAAAAGCAATCTTATATCTGGCTCAGTAATGTACATCGAGGAGAAAACAAAGACC
AAGTACACAGGAAATCCAACAAAGATGTATGAAGTGGTCTACCAAATTGGAACAGAGACTCG
TTCAGACTTCTATGACATCGTCTTGGTGGCCACTCCGTTGAATCGAAAAATGTGCAATATTA
CTTTTCTCAACTTTGATCCTCCAATTGAGGAATTCCATCAATATTATCAACATATAGTGACA
ACTTTAGTTAAGGGGGAATTGAATACATCTATCTTTAGCTCTAGACCCATAGATAAATTTGG
CCTTAATACAGTTTTTAACCACTGATAATTGAGATTTGTTTCATTAAACAGTATTGGGATTGTGC
CCTCTGTGAGAGAAAAGGAAGATCCTGAGCCATCAACAGATGGAACATATGTTTGGAAGATC
TTTTCCCAAGAACTCTTACTAAAGCACAAATTTTAAAGCTCTTCTGTCCTATGATTATGC
TGTGAAGAAGCCATGGCTTGCAATATCCTCACTATAAGCCCCCGGAGAAATGCCCTCTATCA
TTCTCCATGATCGACTTTATTACCTCAATGGCATAGAGTGTGCAGCAAGTGCCATGGAGATG
AGTGCCATTGCAGCCCACAACGCTGCACTCCTTGCCATATCACCGCTGGAACGGGCACACAGA
CATGATTGATCAGGATGGCTTATATGAGAACTTAAACTGAACTAT**TGA**AGTGACACACTCC
TTTTTCCCCTCCTAGTTCCAAATGACTATCAGTGGCAAAAAAGAACAAAATCTGAGCAGAGA
TGATTTTGAACCAGATATTTTGCCATTATCATTTGTTTAATAAAAGTAATCCCTGCTGGTCAT
AGGAAAAAAAAAAAAA

FIGURE 26

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62880

<subunit 1 of 1, 505 aa, 1 stop

<MW: 56640, pI: 6.10, NX(S/T): 4

MGRVVAELVSSLLGLWLLLCSCGCPEGAE LRAPPDKIAIIGAGIGGTSAAYYLRQKFGKDVK
IDLFEREEVGGRLATMMVQGGQEYEAGGSVIHPLNLHMKRFVKDLGLSAVQASGGLLGIYNGE
TLVFEE SNWFIINVIKLVWRYGFQSLRMH MWVEDVLDKFMRIYRYQSHDYAFSSVEKLLHAL
GGDDFLGMLNRTLLET LQKAGFSEKFLNEMIAPVMRVNYGQSTDINAFVGAVSLSCSDSGLW
AVEGGNKLVC SGLLQASKSNLISGSVMYIEEKT KTKYTGNPTKMYEVVYQIGTETRSDFYDI
VLVATPLNRKMSNITFLNFDPPIEEFHQYYQHIVTTLVKGELNTSIFSSRPIDKFG LNTVLT
TDNSDLFINSIGIVPSVREKEDPEPSTDGTYVWKIFSQETLTKAQILKLFLSYDYAVKKPWL
AYPHYKPPKEKCP SIILHDRLYYLN GIECAASAMEMSAIAAHNAALLAYHRWNGHTDMIDQDG
LYEKLKTEL

Important features:**Signal peptide:**

amino acids 1-23

N-glycosylation sites.

amino acids 196-200, 323-327, 353-357

Tyrosine kinase phosphorylation site.

amino acids 291-298

N-myristoylation sites.

amino acids 23-29, 41-47, 43-49, 45-51, 46-52, 72-78, 115-121,
119-125, 260-266, 384-390, 459-465

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 12-23, 232-243

27/75

FIGURE 27

CATTTCCAACAAGAGCACTGGCCAAGTCAGCTTCTTCTGAGAGAGTCTCTAGAAGAC**ATG**AT
GCTACACTCAGCTTTGGGTCTCTGCCTCTTACTCGTCACAGTTTCTTCCAACCTTGCCATTG
CAATAAAAAAGGAAAAGAGGCCTCCTCAGACACTCTCAAGAGGATGGGGAGATGACATCACT
TGGGTACAACTTATGAAGAAGGTCTCTTTTATGCTCAAAAAAGTAAGAAGCCATTAATGGT
TATTCATCACCTGGAGGATTGTCAATACTCTCAAGCACTAAAGAAAGTATTTGCCCAAATG
AAGAAATACAAGAAATGGCTCAGAATAAGTTCATCATGCTAAACCTTATGCATGAAACCACT
GATAAGAATTTATCACCTGATGGGCAATATGTGCCTAGAATCATGTTTGTAGACCCTTCTTT
AACAGTTAGAGCTGACATAGCTGGAAGATACTCTAACAGATTGTACACATATGAGCCTCGGG
ATTTACCCCTATTGATAGAAAACATGAAGAAAGCATTAAAGACTTATTCAGTCAGAGCTA**TAA**
GAGATGATGGAAAAAAGCCTTCACTTCAAAGAAGTCAAATTTTCATGAAGAAAACCTCTGGCA
CATTGACAAATACTAAATGTGCAAGTATATAGATTTTGTAATATTACTATTTAGTTTTTTTA
ATGTGTTTGCAATAGTCTTATTAAATAAATGTTTTTTAAATCTGA

28/75

FIGURE 28

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64896

<subunit 1 of 1, 166 aa, 1 stop

<MW: 19171, pI: 8.26, NX(S/T): 1

MMLHSALGLCLLLVTVSSNLAIKKEKRPPQTLSRGWGDDITWVQTYEEGLFYAQKSKKPL
MVIHHLEDCQYSQALKKVFAQNEEIQEMAQNKFIMLNLMHETTDKNLSPDGQYVPRIMFVDP
SLTVRADIAGRYSNRLYTYEPRDLPLLIENMKKALRLIQSEL

Important features:

Signal peptide:

amino acids 1-23

N-myristoylation site.

amino acids 51-57

29/75

FIGURE 29

TAAAACAGCTACAATATTCCAGGGCCAGTCACTTGCCATTTCTCATAACAGCGTCAGAGAGA
AAGAACTGACTGAAACGTTTGAG**ATG**AAGAAAGTTCTCCTCCTGATCACAGCCATCTTGGCA
GTGGCTGTTGGTTTCCCAGTCTCTCAAGACCAGGAACGAGAAAAAAGAAGTATCAGTGACAG
CGATGAATTAGCTTCAGGGTTTTTTGTGTTCCCTTACCCATATCCATTTTCGCCCACTTCCAC
CAATTCCATTTCCAAGATTTCCATGGTTTAGACGTAATTTTCTATTCCAATACCTGAATCT
GCCCCTACAACCTCCCCTTCCTAGCGAAAAG**TAA**ACAAGAAGGATAAGTCACGATAAACCTGG
TCACCTGAAATTGAAATTGAGCCACTTCCTTGAAGAATCAAAATTCCTGTTAATAAAAGAAA
AACAAATGTAATTGAAATAGCACACAGCATTCTCTAGTCAATATCTTTAGTGATCTTCTTTA
ATAAACATGAAAGCAAAGATTTTGGTTTCTTAATTTCCACA

30/75

FIGURE 30

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71290

><subunit 1 of 1, 85 aa, 1 stop

><MW: 9700, pI: 9.55, NX(S/T): 0

MKKVLLLLITAILAVAVGFPVSQDQEREKRSISDSDELASGFFVFPYPYPFRPLPPIPFPRFP
WFRRNFPIPIPIESAPTTPLPSEK

Important features of the protein:

Signal peptide:

amino acids 1-17

Homologous region to B3-hordein:

amino acids 47-85

31/75

FIGURE 31

CGGACGCGTGGGCGGGCGCGCCGGGAGGGACCGGCGGCGGC**ATG**GGCGGGGGCCCTGGGAT
GCGGGCCCGTCTCGCCGCTGCTGCCGCTGTTGCTGCTGCTCGGCCTGGCCCGGGCGCCGC
GGGAGCGCCGGGCCCCGACGGTTTAGACGTCTGTGCCACTTGCCATGAACATGCCACATGCC
AGCAAAGAGAAGGGAAGAAGATCTGTATTTGCAACTATGGATTTGTAGGGAACGGGAGGACT
CAGTGTGTTGATAAAAATGAGTGCCAGTTTGGAGCCACTCTTGTCTGTGGGAACACACATC
TTGCCACAACACCCCCGGGGGCTTCTATTGCATTTGCCTGGAAGGATATCGAGCCACAAACA
ACAACAAGACATTTCATTCCCAACGATGGCACCTTTTGTACAGACATAGATGAGTGTGAAGTT
TCTGGCCTGTGCAGGCATGGAGGGCGATGCGTGAACACTCATGGGAGCTTTGAATGCTACTG
TATGGATGGATACTTGCCAAGGAATGGACCTGAACCTTTCCACCCGACCACCGATGCCACAT
CATGCACAGAAATAGACTGTGGTACCCCTCCTGAGGTTCCAGATGGCTATATCATAGGAAAT
TATACGTCTAGTCTGGGCAGCCAGGTTTCGTTATGCTTGCAGAGAAGGATTTCTCAGTGTTCC
AGAAGATACAGTTTCAAGCTGCACAGGCGCTGGGCACATGGGAGTCCCCAAAATTACATTGCC
AAGAGATCAACTGTGGCAACCCTCCAGAAATGCGGCACGCCATCTTGTTAGGAAATCACAGC
TCCAGGCTGGGCGGTGTGGCTCGCTATGTCTGTCAAGAGGGCTTTGAGAGCCCTGGAGGAAA
GATCACTTCTGTTTGCACAGAGAAAGGCACCTGGAGAGAAAGTACTTTAACATGCACAGAAA
TTCTGACAAAGATTAATGATGTATCACTGTTTAAATGATACCTGTGTGAGATGGCAAATAAAC
TCAAGAAGAATAAAACCCCAAGATCTCATATGTGATATCCATAAAAGGACAACGGTTGGACCC
TATGGAATCAGTTTCGTGAGGAGACAGTCAACTTGACCACAGACAGCAGGACCCCAAGTGT
GCCTAGCCCTGTACCCAGGCACCAACTACACCGTGAACATCTCCACAGCACCTCCCAGGCGC
TCGATGCCAGCCGTCATCGGTTTCCAGACAGCTGAAGTTGATCTCTTAGAAGATGATGGAAG
TTTCAATATTTCAATATTTAATGAACTTGTGTTGAAATTGAACAGGCGTTCTAGGAAAGTTG
GATCAGAACACATGTACCAATTTACCGTTCTGGGTGAGAGGTGGTATCTGGCTAACTTTTCT
CATGCAACATCGTTTAACTTCACAACGAGGGAACAAGTGCCTGTAGTGTGTTTGGATCTGTA
CCCTACGACTGATTATACGGTGAATGTGACCCTGCTGAGATCTCCTAAGCGGCACCTCAGTGC
AAATAACAATAGCAACTCCCCCAGCAGTAAACAGACCATCAGTAACATTTTCAGGATTTAAT
GAAACCTGCTTGAGATGGAGAAGCATCAAGACAGCTGATATGGAGGAGATGTATTTATTCCA
CATTTGGGGCCAGAGATGGTATCAGAAGGAATTTGCCCAGGAAATGACCTTTAATATCAGTA
GCAGCAGCCGAGATCCCGAGGTGTGCTTGGACCTACGTCCGGGTACCAACTACAATGTCAGT
CTCCGGGCTCTGTCTTCGGAACCTTCTGTGGTCACTCTCCCTGACAACCCAGATAACAGAGCC
TCCCCTCCCGGAAGTAGAATTTTTTACGGTGCACAGAGGACCTCTACCACGCCTCAGACTGA
GGAAAGCCAAGGAGAAAAATGGACCAATCAGTTCATATCAGGTGTTAGTGCTTCCCCTGGCC
CTCCAAAGCACATTTTCTTGTGATTCTGAAGGCGCTTCCCTCCTTCTTTAGCAACGCCTCTGA
TGCTGATGGATACGTGGCTGCAGAACTACTGGCCAAAGATGTTCCAGATGATGCCATGGAGA
TACCTATAGGAGACAGGCTGTACTATGGGGAATATTATAATGCACCCCTTGAAAAGAGGGAGT
GATTACTGCATTATATTACGAATCACAAGTGAATGGAATAAGGTGAGAAGACACTCCTGTGC
AGTTTGGGCTCAGGTGAAAGATTCGTCACTCATGCTGCTGCAGATGGCGGGTGTTGGACTGG
GTTCCCTGGCTGTTGTGATCATTCTCACATTCCTCTCCTTCTCAGCGGTG**TGA**TGGCAGATG
GACACTGAGTGGGGAGGATGCACTGCTGCTGGGCAGGTGTTCTGGCAGCTTCTCAGGTGCC
GCACAGAGGCTCCGTGTGACTTCCGTCCAGGGAGCATGTGGGCCTGCAACTTTCTCCATTCC
CAGCTGGGCCCCATTCTGGATTTAAGATGGTGGCTATCCCTGAGGAGTCACCATAAGGAGA
AACTCAGGAATTCTGAGTCTTCCCTGCTACAGGACCAGTTCTGTGCAATGAACTTGAGACT
CCTGATGTACACTGTGATATTGACCGAAGGCTACATACAGATCTGTGAATCTTGGCTGGGAC
TTCTCTGAGTGTGCTGAGGGTCAGCTCCTCTAGACATTGACTGCAAGAGAATCTCTGCA
ACCTCCTATATAAAAGCATTTCTGTTAATTCATTTCAGAAATCCATTCTTTACAATATGCAGTG
AGATGGGCTTAAGTTTGGGCTAGAGTTTGACTTTATGAAGGAGGTCAATTGAAAAAGAGAAC
GTGACGTAGGCAAATGTTTCAAGCACTTTAGAAACAGTACTTTTCCCTATAATTAGTTGATAT
ACTAATGAGAAAATATACTAGCCTGGCCATGCCAATAAGTTTCCCTGCTGTGTCTGTTAGGCA
GCATTGCTTTGATGCAATTTCTATTGTCCTATATATTCAAAGTAATGTCTACATTCCAGTA
AAAATATCCCGTAATTAAAAA

32/75

FIGURE 32

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96031
><subunit 1 of 1, 747 aa, 1 stop
><MW: 82710, pI: 6.36, NX(S/T): 18
MGRGPWDAGPSRRLPLLLLLGLARGAAGAPGPDGLDVCATCHEHATCQOREGKKICICNYG
FVGNGRTQCVDKNECQFGATLVCGNHTSCHNTPGGFYCICLEGYRATNNNKTFIPNDGTFCT
DIDECEVSGLCRHGGRCVNTHGSEFCYCMDGYLPRNGPEPFHPTTDATEIDCGTPPEVP
DGYIIGNYTSSLGSQVRYACREGFFSVPEDTVSSCTGLGTWESPCLHCQEINCGNPPEMRHA
ILVGNHSSRLGGVARYVCQEGFESPGGKITSVCTEKGWRESTLTCTEILTAKINDVSLFNDT
CVRWQINSRRINPKISYVISIKGQRLDPMESVREETVNLTTDSRTPEVCLALYPGTNYTVNI
STAPPRRSMFAVIGFQTAEDVLEDDGSFNISIFNETCLKLNRRSRKVGSEHMYQFTVLGQR
WYLANFSHATSFNFTTREQVPVVCCLDLYPTTDYTVNVTLRSPKRHSVQITATPPAVKQTI
SNISGFNETCLRWSIKTADMEEMYLFIHWGQRWYQKEFAQEMTFNIISSSSRDPEVCLDLRP
GTNYNVSLRALSSELPVVISLTQTITEPPLPEVEFFTVHRGPLPRLRLRKAKEKNGPISSYQ
VLVLPLALQSTFSCDSEGASSFFSNASDADGYVAAELLAKDVPDDAMEIPIGDRLYYGEYYN
APLKRGSYCIILRITSEWNKVRHSCAVWAQVKDSSLMLLQMAVGVLGSLAVVILTFLSF
SAV
```

Important features of the protein:

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 718-740

N-glycosylation sites.

amino acids 87-91, 112-116, 193-197, 253-257, 308-312, 348-352, 367-371, 371-375, 402-406, 407-411, 439-443, 447-451, 470-474, 498-502, 503-507, 542-546, 563-567, 645-649

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 478-482, 686-690, 705-709

Tyrosine kinase phosphorylation site.

amino acids 419-427

N-myristoylation sites.

amino acids 22-28, 35-41, 65-71, 86-92, 96-102, 120-126, 146-152, 192-198, 252-258, 274-280, 365-371, 559-565, 688-694, 727-733.

Amidation site.

amino acids 52-56

Aspartic acid and asparagine hydroxylation sites.

amino acids 91-103, 141-153.

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 624-635

Cytochrome c family heme-binding site signature.

amino acids 39-45

Calcium-binding EGF-like domain proteins pattern proteins.

amino acids 85-106, 135-156

Receptor tyrosine kinase class V proteins:

amino acids 389-422

FIGURE 33

GGAAAAGGTACCCGCGAGAGACAGCCAGCAGTTCTGTGGAGCAGCGGTGGCCGGCTAGG**ATG**
GGCTGTCTCTGGGGTCTGGCTCTGCCCCTTTTCTTCTTCTGCTGGGAGGTTGGGGTCTCTGG
GAGCTCTGCAGGCCCCAGCACCCGCGAGAGCAGACACTGCGATGACAACGGACGACACAGAAG
TGCCCGCTATGACTCTAGCACCCGGGCCACGCCGCTCTGGAAACTCAAACGCTGAGCGCTGAG
ACCTCTTCTAGGGCCTCAACCCAGCCGGCCCCATTCCAGAAGCAGAGACCAGGGGAGCCAA
GAGAATTTCCCCTGCAAGAGAGACCAGGAGTTTCACAAAACATCTCCCAACTTCATGGTG
TGATCGCCACCTCCGTGGAGACATCAGCCGCCAGTGGCAGCCCCGAGGGAGCTGGAATGACC
ACAGTTCAGACCATCACAGGCAGTGATCCCGAGGAAGCCATCTTTGACACCCTTTGCACCGA
TGACAGCTCTGAAGAGGCAAAGACACTCACAATGGACATATTGACATTGGCTCACACCTCCA
CAGAAGCTAAGGGCCTGTCCTCAGAGAGCAGTGCCCTCTCCGACGGCCCCCATCCAGTCATC
ACCCCGTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCC
GTCACGGGCCTCAGAGAGCAGCGCCTCTTCCGACGGCCCCCATCCAGTCATCACCCCGTCAT
GGTCCCCGGGATCTGATGTCACTCTCCTCGCTGAAGCCCTGGTGACTGTACAAACATCGAG
GTTATTAATTGCAGCATCACAGAAATAGAAACAACAACTTCCAGCATCCCTGGGGCCTCAGA
CATAGATCTCATCCCCACGGAAGGGGTGAAGGCCTCGTCCACCTCCGATCCACCAGCTCTGC
CTGACTCCACTGAAGCAAAACCACACATCACTGAGGTCACAGCCTCTGCCGAGACCCTGTCC
ACAGCCGGCACCACAGAGTCAGCTGCACCTCATGCCACGGTTGGGACCCCACTCCCCACTAA
CAGCGCCACAGAAAGAGAAGTGACAGCACCCGGGGCCACGACCCTCAGTGGAGCTCTGGTCA
CAGTTAGCAGGAATCCCCTGGAAGAAACCTCAGCCCTCTCTGTTGAGACACCAAGTTACGTC
AAAGTCTCAGGAGCAGCTCCGGTCTCCATAGAGGCTGGGTCAGCAGTGGGCAAAACAACCTTC
CTTTGCTGGGAGCTCTGCTTCCTCCTACAGCCCTCGGAAGCCGCCCTCAAGAACTTCACCC
CTTCAGAGACACCGACCATGGACATCGCAACCAAGGGGCCCTTCCCCACCAGCAGGGACCCT
CTTCCTTCTGTCCCTCCGACTACAACCAACAGCAGCCGAGGGACGAACAGCACCTTAGCCAA
GATCACAACCTCAGCGAAGACCACGATGAAGCCCCAACAGCCACGCCCACGACTGCCCGGAC
GAGGCCGACCACAGACG**TGAG**TGCAGGTGAAAATGGAGGTTTCTCCTCCTGCGGCTGAGTG
TGGCTTCCCCGGAAGACCTCACTGACCCAGAGTGGCAGAAAGGCTGATGCAGCAGCTCCAC
CGGGAACCTCACGCCCACGCGCCTCACTTCCAGGTCTCCTTACTGCGTGTGAGGAGAGGCTA
ACGGACATCAGCTGCAGCCAGGCATGTCCCGTATGCCAAAAGAGGGTGCTGCCCCTAGCCTG
GGCCCCCACCGACAGACTGCAGCTGCGTTACTGTGCTGAGAGGTACCCAGAAGGTTCCCATG
AAGGGCAGCATGTCCAAGCCCCTAACCCAGATGTGGCAACAGGACCCTCGCTCACATCCAC
CGGAGTGTATGTATGGGGAGGGGCTTCACCTGTTCCAGAGGTGTCCTTGGACTCACCTTGG
CACATGTTCTGTGTTTCAGTAAAGAGAGACCTGATCACCCATCTGTGTGCTTCCATCCTGCA
TTAAAATTCACTCAGTGTGGCCCCAAAAAAA

34/75

FIGURE 34

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108722
><subunit 1 of 1, 482 aa, 1 stop
><MW: 49060, pI: 4.74, NX(S/T): 4
MGCLWGLALPLFFFCWEVGVSGSSAGPSTRRADTAMTTDDTEVPAMTLAPGHAALETQTLISA
ETSSRASTPAGPIPEAETRGAKRISPARETRSF TKTS PNF MVLIATSVETSAASGSPEGAGM
TTVQTITGSDPEEAIFDTLCTDDSS EEA KTLTMDILT LAHTSTEAKGLSSESSASSDGPHPV
ITPSRASESSASSDGPHPVITPSRASESSASSDGPHPVITPSWSPGSDVTL LAEALVTVTNI
EVINCSITEIETTTSSIPGASDIDLIPTEGVKASSTSDPPALPDSTEAKPHITEVTASAETL
STAGTTESAAPHATVGTPLPTNSATEREVTAPGATTL SGALVTVSRNPLEETSALSVETPSY
VKVSGAAPV SIEAGSAVGKTTSFAGSSASSYSPSEAALKNFTPSETPTMDIATKGPFPTSRD
PLPSVPPTTTN SSRGTNSTLAKITTS AKTTMKPQQPRPRLPGRGRPQT
```

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation sites.

amino acids 252-256, 445-449, 451-455

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 84-88

N-myristoylation sites.

amino acids 2-8, 19-25, 117-123, 121-127, 232-238, 278-284,
314-320, 349-355, 386-392, 397-403, 449-455

ATP/GTP-binding site motif A (P-loop).

amino acids 385-393

FIGURE 35

GCCTCTGAATTGTTGGGCAGTCTGGCAGTGGAGCTCTCCCCGGTCTGACAGCCACTCCAGAG
GCC**ATG**CTTCGTTTCTTGCCAGATTTGGCTTTCAGCTTCCTGTTAATTCTGGCTTTGGGCCA
GGCAGTCCAATTTCAAGAATATGTCTTTCTCCAATTTCTGGGCTTAGATAAGGCGCCTTCAC
CCCAGAAGTTCCAACCTGTGCCTTATATCTTGAAGAAAATTTTCCAGGATCGCGAGGCAGCA
GCGACCACTGGGGTCTCCCGAGACTTATGCTACGTAAAGGAGCTGGGCGTCCGCGGGAATGT
ACTTCGCTTTCTCCAGACCAAGGTTTCTTTCTTTACCCAAAGAAAATTTCCCAAGCTTCCT
CCTGCCTGCAGAAGCTCCTCTACTTTAACCTGTCTGCCATCAAAGAAAGGGAACAGTTGACA
TTGGCCAGCTGGGCCTGGACTTGGGGCCCAATTCTTACTATAACCTGGGACCAGAGCTGGA
ACTGGCTCTGTTCTGTTTCAGGAGCCTCATGTGTGGGGCCAGACCACCCCTAAGCCAGGTA
AAATGTTTGTGTTGCGGTCAGTCCCATGGCCACAAGGTGCTGTTCACTTCAACCTGCTGGAT
GTAGCTAAGGATTGGAATGACAACCCCCGGAAAAATTTCTGGGTTATTCCTGGAGATACTGGT
CAAAGAAGATAGAGACTCAGGGGTGAATTTTCAGCCTGAAGACACCTGTGCCAGACTAAGAT
GCTCCCTTCATGCTTCCCTGCTGGTGGTGA CTCTCAACCCTGATCAGTGCCACCCTTCTCGG
AAAAGGAGAGCAGCCATCCCTGTCCCCAAGCTTTCTTGTAAGAACCTCTGCCACCGTCACCA
GCTATTCATTA ACTTCCGGGACCTGGGTTGGCACAAGTGGATCATTGCCCCCAAGGGGTTCA
TGGCAAATTACTGCCATGGAGAGTGTCCCTTCTCACTGACCATCTCTCTCAACAGCTCCAAT
TATGCTTTCATGCAAGCCCTGATGCATGCCGTTGACCCAGAGATCCCCCAGGCTGTGTGTAT
CCCCACCAAGCTGTCTCCCATTTCCATGCTCTACCAGGACAATAATGACAATGTCATTCTAC
GACATTATGAAGACATGGTAGTCGATGAATGTGGGTGTGGG**TAG**GATGTCAGAAATGGGAAT
AGAAGGAGTGTTCCTTAGGGTAAATCTTTTAATAAAACTACCTATCTGGTTTATGACCACTTA
GATCGAAATGTC

FIGURE 36

MLRFLPDLAFSFLILALGQAVQFQEYVFLQFLGLDKAPSPQKFQVPYILKKIFQDREAAA
TTGVSRDLCYVKELGVRGNVLRFLPDQGFFLYPKKISQASSCLQKLLYFNLSAIKEREQLTL
AQLGLDLGPNSYYNLGPELELALFLVQEPHVWGQTPKPGKMFVLRSPWPQGAVHFNLLDV
AKDWNDNPRKNFGLFLEILVKEDRD SGVNFQPEDTCARLRCSLHASLLVVTLNPDQCHPSRK
RRAAIPVPKLSCKNLCHRHQLFINFRDLGWHKWIIAPKGFMANYPCHGECPFSLTISLNSSNY
AFMQALMHAVDPEIPQAVCIPTKLSPISMLYQDNNDNVILRHYEDMVVDECGCG

Important features of the protein:**Signal peptide:**

amino acids 1-21

N-glycosylation sites.

amino acids 112-116, 306-310

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 96-100

N-myristoylation site.

amino acids 77-83

TGF-beta family proteins.

amino acids 264-299, 327-341, 345-364

37/75

FIGURE 37

CAC TTTCTCCCTCTCTTCTTTACTTTTCGAGAAACCGCGCTTCCGCTTCTGGTCGCAGAGAC
CTCGGAGACCGCGCCGGGGAGACGGAGGTGCTGTGGGTGGGGGGGACCTGTGGCTGCTCGTA
CCGCCCCCACCCTCCTCTTCTGCACTGCCGTCTCCGGAAGACCTTTTCCCCTGCTCTGTT
TCCTTCACCGAGTCTGTGCATCGCCCCGGACCTGGCCGGGAGGAGGCTTGGCCGGCGGGAGA
TGCTCTAGGGGCGGCGCGGGAGGAGCGGCCGGGACGGAGGGCCCCGGCAGGAAG**ATG**GGC
TCCCGTGGACAGGGACTCTTGCTGGCGTACTGCCTGCTCCTTGCCTTTGCCTCTGGCCTGGT
CCTGAGTCGTGTGCCCCATGTCCAGGGGGAACAGCAGGAGTGGGAGGGGACTGAGGAGCTGC
CGTCGCCTCCGGACCATGCCGAGAGGGCTGAAGAACAACATGAAAAATACAGGCCCAGTCAG
GACCAGGGGCTCCCTGCTTCCCGGTGCTTGCCTGCTGTGACCCCGGTACCTCCATGTACCC
GGCGACCGCCGTGCCCCAGATCAACATCACTATCTTGAAAGGGGAGAAGGGTGACCGCGGAG
ATCGAGGCTCCAAGGGAAATATGGCAAAACAGGCTCAGCAGGGGGCCAGGGGCCACACTGGA
CCCAAAGGGCAGAAGGGCTCCATGGGGGCCCTTGGGAGCGGTGCAAGAGCCACTACGCCGC
CTTTTCGGTGGGCGGAAGAAGCCATGCACAGCAACCACTACTACAGACGGTGATCTTCG
ACACGAGTTCGTGAACCTCTACGACCACTTCAACATGTTACCGGCAAGTTCTACTGCTAC
GTGCCCCGCCTCTACTTCTTACGCCTCAACGTGCACACCTGGAACCAGAAGGAGACCTACCT
GCACATCATGAAGAACGAGGAGGAGGTGGTGATCTTGTTTCGCGCAGGTGGGCGACCGCAGCA
TCATGCAAAGCCAGAGCCTGATGCTGGAGCTGCGAGAGCAGGACCAGGTGTGGGTACGCCTC
TACAAGGGCGAACGTGAGAACGCCATCTTACGCGAGGAGCTGGACACCTACATCACCTTCAG
TGGCTACCTGGTCAAGCACGCCACCGAGCCC**TAG**CTGGCCGGCCACCTCCTTTCTCTCGCC
ACCTTCCACCCCTGCGCTGTGCTGACCCCAACGCTCTTCCCCGATCCCTGGACTCCGACTC
CCTGGCTTTGGCATTCAAGTGAAGACGCCCTGCACACACAGAAAGCCAAAGCGATCGGTGCTCC
CAGATCCCGCAGCCTCTGGAGAGAGCTGACGGCAGATGAAATCACCAGGGCGGGGCACCCGC
GAGAACCCTCTGGGACCTTCCGCGGCCCTCTCTGCACACATCCTCAAGTGACCCCGCACGGC
GAGACGCGGGTGGCGGCAGGGCGTCCCAGGGTGGCGCACCGCGGCTCCAGTCTTTGGAATA
ATTAGGCAAATTCTAAAGGTCTCAAAGGAGCAAAGTAAACCGTGGAGGACAAAGAAAAGGG
TTGTTATTTTTGTCTTTCCAGCCAGCCTGCTGGCTCCCAAGAGAGAGGCCTTTTCAGTTGAG
ACTCTGCTTAAGAGAAGATCCAAAGTTAAAGCTCTGGGGTCAGGGGAGGGGGCCGGGGGCAGG
AAACTACCTCTGGCTTAATTCTTTTAAGCCACGTAGGAACCTTTCTTGAGGGATAGGTGGACC
CTGACATCCCTGTGGCCTTGCCCAAGGGCTCTGCTGGTCTTTCTGAGTCACAGCTGCGAGGT
GATGGGGGCTGGGGCCCCAGGCGTCAGCCTCCCAGAGGGACAGCTGAGCCCCCTGCCTTGGC
TCCAGGTTGGTAGAAGCAGCCGAAGGGCTCCTGACAGTGGCCAGGGACCCCTGGGTCCCCCA
GGCCTGCAGATGTTTCTATGAGGGGCAGAGCTCCTTGGTACATCCATGTGTGGCTCTGCTCC
ACCCCTGTGCCACCCCAAGAGCCCTGGGGGGTGGTCTCCATGCCTGCCACCTGGCATCGGCT
TTCTGTGCCGCCTCCACACAAATCAGCCCCAGAAGGCCCGGGGCCTTGGCTTCTGTTTTT
TATAAAACACCTCAAGCAGCACTGCAGTCTCCATCTCCTCGTGGGCTAAGCATCACCGCTT
CCACGTGTGTTGTGTTGGTTGGCAGCAAGGCTGATCCAGACCCCTTCTGCCCCACTGCCCT
CATCCAGGCCTCTGACCAGTAGCCTGAGAGGGGCTTTTTCTAGGCTTCAGAGCAGGGGAGAG
CTGGAAGGGGCTAGAAAGCTCCCGCTTGTCTGTTTCTCAGGCTCCTGTGAGCCTCAGTCCTG
AGACCAGAGTCAAGAGGAAGTACACGTCCCAATCACCCGTGTCAGGATTCAGTCTCAGGAGC
TGGGTGGCAGGAGAGGCAATAGCCCCCTGTGGCAATTGCAGGACCAGCTGGAGCAGGGTTGCG
GTGTCTCCACGGTGCTCTCGCCCTGCCCATGGCCACCCAGACTCTGATCTCCAGGAACCC
ATAGCCCCCTCTCCACCTCACCCCATGTTGATGCCCAGGGTCACTCTTGCTACCCGCTGGGCC
CCCAAACCCCGCTGCCTCTCTTCTTCCCCCATCCCCACCTGGTTTTGACTAATCCTGC
TTCCCTCTCTGGGCCTGGCTGCCGGGATCTGGGGTCCCTAAGTCCCTCTCTTTAAAGAACTT
CTGCGGGTCAAGTCTGAAGCCGAGTTGCTGTGGGCGTGCCCGGAAGCAGAGCGCCACACTC
GCTGCTTAAGCTCCCCCAGCTCTTTCCAGAAAACATTAAACTCAGAATTGTGTTTTCAA

38/75

FIGURE 38

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41234
><subunit 1 of 1, 281 aa, 1 stop
><MW: 31743, pI: 6.83, NX(S/T): 1
MGSRGQGLLLAYCLLLAFASGLVLSRVPHVQGEQQEWEGTEELPSPPDHAERAEEQHEKYRP
SQDQGLPASRCLRCCDPGTSMPATAVPQINITILKGEKGDRGDRGLQGKYGKTGSAGARGH
TGPKGQKGSMGAPGERCKSHYAAFSVGRKKPMHSNHYYQTVIFDTEFVNLYDHFNMFTGKFY
CYVPGLYFFSLNVHTWNQKETYLHIMKNEEEVVILFAQVGDRSIMQSQSLMLELREQDQVWV
RLYKGERENAI FSEELDTYITFSGYLVKHATEP
```

Signal sequence.

amino acids 1-25

N-glycosylation site.

amino acids 93-97

N-myristoylation sites.

amino acids 7-13, 21-27, 67-73, 117-123, 129-135

Amidation site.

amino acids 150-154

Cell attachment sequence.

amino acids 104-107

FIGURE 39

GAATTCGGCACGAGGGAAGAAGAGAAAGAAAATCTCCGGGGCTGCTGGGAGCATATAAAGAA
GCCCTGTGGCCTTGCTGGTTTTACCATCCAGACCAGAGTCAGGCCACAGACGGAC**ATGG**CTG
CTCAAGGCTGGTCCATGCTCCTGCTGGCTGTCCTTAACCTAGGCATCTTCGTCCGTCCCTGT
GACACTCAAGAGCTACGATGTCTGTGTATTCAGGAACACTCTGAATTCATTCCTCTCAAAC
CATTA AAAATATAATGGTGATATTCGAGACCATTTACTGCAACAGAAAGGAAGTGATAGCAG
TCCCAAAAATGGGAGTATGATTTGTTTGGATCCTGATGCTCCATGGGTGAAGGCTACTGTT
GGCCAATTACTAACAGGTTCCCTACCTGAGGACCTCAAACAAAAGGAATTTCCACCGGCAAT
GAAGCTTCTGTATAGTGTTGAGCATGAAAAGCCTCTATATCTTTCATTTGGGAGACCTGAGA
ACAAGAGAATATTTCCCTTTCCAATTCGGGAGACCTCTAGACACTTTGCTGATTTAGCTCAC
AACAGTGATAGGAATTTTCTACGGGACTCCAGTGAAGTCAGCTTGACAGGCAGTGATGCC**TA**
AAAGCCACTCATGAGGCAAAGAGTTTCAAGGAAGCTCTCCTCCTGGAGTTTTGGCGTTCTCA
TTCTTATACTCTATTCCCGCGTTAGTCTGGTGTATGGATCTATGAGCTCTCTTTTAATATTT
TATTATAAATGTTTTATTTACTTAACTTCCTAGTGAATGTTACAGGTGACTGCTCCCCCAT
CCCCATTTCTTGATATTACATATAATGGCATCATATAACCCCTTTATTGACTGACAAACTACT
CAGATTGCTTAACATTTTGTGCTTCAAAGTCTTATCCCACTCCACTATGGGCTGTTACAGAG
TGCATCTCGGTGTAGAGCAAGGCTCCTTGTCTTCAGTGCCCCAGGGTGAAATACTTCTTTGA
AAAATTTTCATTCATCAGAAAATCTGAAATAAAAATATGTCTTAATTGAG

40/75

FIGURE 40

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73838
><subunit 1 of 1, 167 aa, 1 stop
><MW: 19091, pI: 7.48, NX(S/T): 1
MAAQGWSMLLLAVLNLGIFVRPCDTQELRCLCIQEHSEFIPLKLIK NIMVIFETIYCNRKEV
IAVPKNGSMICLDPDAPWVKATVGPITNRFLPEDLKQKEFPPAMKLLYSVEHEKPLYLSFGR
PENKRIFPFPIRETSRHFADLAHNSDRNFLRDSSEVSLTGSDA
```

Important features of the protein:

Signal peptide:

amino acids 1-25

N-glycosylation site.

amino acids 68-72

N-myristoylation site.

amino acids 69-75

Small cytokines (intercrine/chemokine) C-x-C subfamily signature

amino acids 40-85

41/75

FIGURE 41

CAGACATGGCTCAGTCACTGGCTCTGAGCCTCCTTATCCTGGTTCTGGCCTTTGGCATCCCC
 AGGACCCAAGGCAGTGATGGAGGGGCTCAGGACTGTTGCCTCAAGTACAGCCAAAGGAAGAT
 TCCCGCCAAGGTTGTCCGCAGCTACCGGAAGCAGGAACCAAGCTTAGGCTGCTCCATCCCAG
 CTATCCTGTTCTTGCCCCGCAAGCGCTCTCAGGCAGAGCTATGTGCAGACCCAAAGGAGCTC
 TGGGTGCAGCAGCTGATGCAGCATCTGGACAAGACACCATCCCCACAGAAACCAGCCCAGGG
 CTGCAGGAAGGACAGGGGGGCCTCCAAGACTGGCAAGAAAGGAAAGGGCTCCAAAGGCTGCA
 AGAGGACTGAGCGGTCACAGACCCCTAAAGGGCCATAGCCCAGTGAGCAGCCTGGAGCCCTG
 GAGACCCCAACCAGCCTCACCAGCGCTTGAAGCCTGAACCCAAGATGCAAGAAGGAGGCTATG
 CTCAGGGGGCCCTGGAGCAGCCACCCCATGCTGGCCTTGCCACACTCTTTCTCCTGCTTTAAC
 CACCCCATCTGCATTCCCAGCTCTACCCTGCATGGCTGAGCTGCCCACAGCAGGCCAGGTCC
 AGAGAGACCGAGGAGGGAGAGTCTCCCAGGGAGCATGAGAGGAGGCAGCAGGACTGTCCCTT
 TGAAGGAGAATCATCAGGACCCTGGACCTGATACGGCTCCCCAGTACACCCACCTCTTCCT
 TGTAATATGATTTATACCTAACTGAATAAAAAGCTGTTCTGTCTTCCCNCCCCA

42/75

FIGURE 42

><MW: 14646, pI: 10.45, NX(S/T): 0

MAQSLALSLILVLAFGIPRTQGS D GGAQDCCLKYSQRKIPAKVVRSYRKQEP
SLGCSIPAILFLPRKRSQAELCADPKELWVQQLMQHLDKTPSPQKPAQGC
RKDRGASKTGKKGKGSKGCKR
TERSQTPKGP

Important features of the protein:

Signal peptide:

amino acids 1-17

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 67-71

N-myristoylation sites.

amino acids 17-23, 23-29, 27-33, 108-114, 118-124, 121-127

Amidation site.

amino acids 112-116

Small cytokines.

amino acids 51-91

43/75

FIGURE 43

AAGGAGCAGCCCGCAAGCACCAAGTGAGAGGC**ATGA**AGTTACAGTGTGTTTCCCTTTGGCTC
 CTGGGTACAATACTGATATTGTGCTCAGTAGACAACCACGGTCTCAGGAGATGTCTGATTTT
 CACAGACATGCACCATATAGAAGAGAGTTTCCAAGAAATCAAAGAGCCATCCAAGCTAAGG
 ACACCTTCCCAAATGTCACCTATCCTGTCCACATTGGAGACTCTGCAGATCATTAAGCCCTTA
 GATGTGTGCTGCGTGACCAAGAACCTCCTGGCGTTCTACGTGGACAGGGTGTTCAGGATCA
 TCAGGAGCCAAACCCCAAATCTTGAGAAAAATCAGCAGCATTGCCAACTCTTTCCCTCTACA
 TGCAGAAAACCTCTGCGGCAATGTCAGGAACAGAGGCAGTGTCACTGCAGGCAGGAAGCCACC
 AATGCCACCAGAGTCATCCATGACAACTATGATCAGCTGGAGGTCCACGCTGCTGCCATTAA
 ATCCCTGGGAGAGCTCGACGTCTTTCTAGCCTGGATTAATAAGAATCATGAAGTAATGTTCT
 CAGCT**TGA**TGACAAGGAACCTGTATAGTGATCCAGGGATGAACACCCCCTGTGCGGTTTACT
 GTGGGAGACAGCCACCTTGAAGGGGAAGGAGATGGGGAAGGCCCTTGCAGCTGAAAGTCC
 CACTGGCTGGCCTCAGGCTGTCTTATTCCGCTTGAAAATAGGCAAAAAGTCTACTGTGGTAT
 TTGTAATAAACTCTATCTGCTGAAAGGGCCTGCAGGCCATCCTGGGAGTAAAGGGCTGCCTT
 CCCATCTAATTTATTGTAAAGTCATATAGTCCATGTCTGTGATGTGAGCCAAGTGATATCCT
 GTAGTACACATTGTACTGAGTGGTTTTTCTGAATAAATTCCATATTTTACCTATGA

44/75

FIGURE 44

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92282
><subunit 1 of 1, 177 aa, 1 stop
><MW: 20452, pI: 8.00, NX(S/T): 2
MKLQCVSLWLLGTILILCSVDNHGLRRCLISTDMHHIEESFQEIKRAIQAKDTFPNVTILST
LETLQIIKPLDVCCVTKNLLAFYVDRVFKDHQEPNPKILRKISSIANSFLYMQKTLRQCQEQ
RQCHCRQEATNATRVIHDNYDQLEVHAAAIKSLGELDVFLAWINKNHEVMFSA
```

Signal sequence:

amino acids 1-18

N-glycosylation sites.

amino acids 56-60, 135-139

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 102-106

N-myristoylation site.

amino acids 24-30

Actinin-type actin-binding domain signature 1.

amino acids 159-169

45/75

FIGURE 45

GCTCCCAGCCAAGAACCTCGGGGCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGCCG
CTAAGCGAGGCCTCCTCCTCCCGCAGATCCGAACGGCCTGGGCGGGGTACCCCGGCTGGGA
CAAGAAGCCGCCGCTGCCTGCCCGGGGCCGGGGAGGGGGCTGGGGCTGGGGCCGGAGGCGG
GGTGTGAGTGGGTGTGTGCGGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG
TGTCTTGGGCACCTACCCGTGGGGCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCGGAG
CCGCCGCGCCGTGAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC
GGCACTCACAGCCCCGCAGCGCATCCCGGTGCGCGCCAGCCTCCCGCACCCCCATCGCCGG
AGCTGCGCCGAGAGCCCCAGGGAGGTGCC**ATG**CGGAGCGGGTGTGTGGTGGTCCACGTATGG
ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGGGCGCCCCCTCGCCTTCTCGGACGCGGGGCC
CCACGTGCACTACGGCTGGGGCGACCCCATCCGCTGCGGCACCTGTACACCTCCGGCCCCC
ACGGGCTCTCCAGCTGCTTCTGCGCATCCGTGCCGACGGCGTCGTGGACTGCGCGCGGGG
CAGAGCGCGCACAGTTTGCTGGAGATCAAGGCAGTGCCTCTGCGGACCGTGGCCATCAAGGG
CGTGACAGCGTGCGGTACCTCTGCATGGGCGCGCAGCGCAAGATGCAGGGGCTGCTTCAGT
ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCAGATGGCTACAATGTGTACCGA
TCCGAGAAGCACCGCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA
CAGAGGCTTTCTTCCACTCTCTCATTTCTGCCATGCTGCCATGGTCCCAGAGGAGCCTG
AGGACCTCAGGGGCCACTTGGAATCTGACATGTTCTCTTCGCCCTGGAGACCGACAGCATG
GACCCATTTGGGCTTGTCACCGGACTGGAGGCCGTGAGGAGTCCAGCTTTGAGAAG**TAACT**
GAGACCATGCCCCGGGCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACG
TGCTTCTACAAGAACAGTCTGAGTCCACGTTCTGTTTAGCTTTAGGAAGAAACATCTAGAA
GTTGTACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTTCTAGCCAATAGACTTGTCTGAT
CATAACATTGTAAGCCTGTAGCTTGCCCAGCTGCTGCCTGGGCCCCCATTTCTGCTCCCTCGA
GGTTGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC
TCACTTCCCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC
CCCAGGAGCAGCCAGAAGACAGGCAGTAGTTTTTAATTTAGGAACAGGTGATCCACTCTGTA
AAACAGCAGGTAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG
ACCATTTGCCCTTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGCATGGCCCACCAG
GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACTTGAGAATTCCCC
CTGAGGCCAGTTCTGTCATGGATGCTGTCCTGAGAATAACTTGCTGTCCCGGTGTCACCTGC
TTCCATCTCCCAGCCCACCAGCCCTCTGCCACCTCACATGCCTCCCCATGGATTGGGGCCT
CCCAGGCCCCCACCCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGAT
TTGAAGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGGAAGACCTAGAAC
CCTTTCCCCAGCACTTGGTTTTCCAACATGATATTTATGAGTAATTTATTTTGATATGTACA
TCTCTTATTTTCTTACATTATTTATGCCCCCAAATTATATTTATGTATGTAAGTGAGGTTTG
TTTTGTATATTAAATGGAGTTTGT

FIGURE 46

MRSGCVVHVHWILAGLWLAVAGRPLAFSDAGPHVHYGWDPIRLRHLYTSGPHGLSSCFLRI
 RADGVVDCARGQSAHSLLEIKAVALLRTVAIKGVHSVRYLCMGADGKMQGLLQYSEEDCAFEE
 EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD
 MFSSPLETDSMDPFGLVTGLEAVRSPSFKEK

Signal peptide:

amino acids 1-22

Casein kinase II phosphorylation site.

amino acids 78-82, 116-120, 190-194, 204-208

N-myristoylation site.

amino acids 15-21, 54-60, 66-72, 201-207

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 48-59

FIGURE 47

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTCAGTGGCCTGATCGCG**ATG**GGGACAAA
GGCGCAAGTCGAGAGGAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG
CATTGGGCAGTGTTACAGTGCACCTCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT
GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCA
AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG
TGACCTTCTTGCCAACCTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC
ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCCTCCTCTGCCACCATTGGGAACCGGG
CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT
GGGATAGTGATGCCTACGAATCCCCAAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCCT
GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCCTGTCAGCCTCTGATACTGGAGAATACA
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT
GTGGAGCGGAATGTGGGGGTCTCGTGGCAGCCGTCCTTGTAACCCTGATTCTCCTGGGAAT
CTTGGTTTTTGGCATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA
CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG
ACCTCGTCATTCCTGGTG**TGA**GCCTGGTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT
CAGGTGCTACCGGACTCTGGCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTC
TACACCCACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC
ATCCTCCTTCATGCCCTCCCTCCCTTTCTTACCCTGCTGAGTGGCCTGGAACCTGTTTTAAA
GTGTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC
TTCTAAGTAGACAGCAAAAATGGCGGGGGTTCGAGGAATCTGCACTCAACTGCCCACCTGGC
TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGTTCTGGGCTCTTTCCTTGTGTACTGAC
GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG
TGATGACACTGGGGTCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG
CCACTGGGATCCCTCTGCCCTGTCTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGT
GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAACTTGAAGCCAAAAG
GATTTAAAACCGCTGCTCTAAAGAAAAGAAAACCTGGAGGCTGGGCGCAGTGGCTCACGCCTG
TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTTCGGGAGTTTCGGGATCAGCCTGACCA
ACATGGAGAAACCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAAACTCCAGCTCAAAAAAAAAAAAAAAAAA

48/75

FIGURE 48

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSLAYSGFSSPRVEW
KFDQGDTTTRLVCYNNKITASYEDRVTFLLPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKV
KLIVLVPPSKPTVNIPSSATIGNRAVLTCSEQDGSPPEYTWFKDGIVMPTNPKSTRAFSNS
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVA AVLVTLI
LLGILVFGIWFAYSRGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

Signal sequence:

amino acids 1-27

Transmembrane domain:

amino acids 238-255

N-glycosylation site.

amino acids 185-189

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 270-274

Casein kinase II phosphorylation site.

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,
193-197, 203-207, 287-291

N-myristoylation site.

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

•

[illegible]

FIGURE 50

MGAARLLPNLTLCLQLLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRI
SATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLGKPSGKSKDCVFTEIVLE
NNYTAFQONARHEGWFMATRQGRPRQASRSRQNQREAHFIKRLYQGQLPFPNHAEKQKQFEF
VGSAPTRRTKRTRRPQPLT

Signal peptide:

amino acids 1-22

N-glycosylation site.

amino acids 9-13, 126-130

cAMP- and cGMP-dependent protein kinase phosphorylation site.

amino acids 60-64

Casein kinase II phosphorylation site.

amino acids 65-69

Tyrosine kinase phosphorylation site.

amino acids 39-48, 89-97

N-myristoylation site.

amino acids 69-75, 188-194

Amidation site.

amino acids 58-62

HBGF/FGF family signature.

amino acids 103-128

FIGURE 51

GTTGTGTCCTTCAGCAAAACAGTGGATTTAAATCTCCTTGCACAAGCTTGAGAGCAACACAA
TCTATCAGGAAAGAAAGAAAGAAAAAACCGAACCTGACAAAAAGAAGAAAAAGAAGA
AAAAAATC**ATG**AAAACCATCCAGCCAAAAATGCACAATTCTATCTCTTGGGCAATCTTCAC
GGGGCTGGCTGCTCTGTGTCTCTTCCAAGGAGTGCCCGTGCGCAGCGGAGATGCCACCTTCC
CCAAAGCTATGGACAACGTGACGGTCCGGCAGGGGGAGAGCGCCACCCTCAGGTGCACTATT
GACAACCGGGTCACCCGGGTGGCCTGGCTAAACCGCAGCACCATCCTCTATGCTGGGAATGA
CAAGTGGTGCCTGGATCCTCGCGTGGTCCTTCTGAGCAACACCCAAACGCAGTACAGCATCG
AGATCCAGAACGTGGATGTGTATGACGAGGGCCCTTACACCTGCTCGGTGCAGACAGACAAC
CACCCAAAGACCTCTAGGGTCCACCTCATTGTGCAAGTATCTCCCAAATTGTAGAGATTTT
TTCAGATATCTCCATTAATGAAGGGAACAATATTAGCCTCACCTGCATAGCAACTGGTAGAC
CAGAGCCTACGGTTACTTGGAGACACATCTCTCCCAAAGCGGTTGGCTTTGTGAGTGAAGAC
GAATACTTGGAATTCAGGGCATCACCCGGGAGCAGTCAGGGGACTACGAGTGCAGTGCCTC
CAATGACGTGGCCGCGCCCGTGGTACGGAGAGTAAAGGTCACCGTGAACCTATCCACCATA
TTTCAGAAGCCAAGGGTACAGGTGTCCCCGTGGGACAAAAGGGGACACTGCAGTGTGAAGCC
TCAGCAGTCCCCTCAGCAGAATTCCAGTGGTACAAGGATGACAAAAGACTGATTGAAGGAAA
GAAAGGGGTGAAAGTGGAAAACAGACCTTTCCTCTCAAACTCATCTTCTTCAATGTCTCTG
AACATGACTATGGGAACCTACACTTGCCTGGCCTCCAACAAGCTGGGCCACACCAATGCCAGC
ATCATGCTATTTGGTCCAGGCGCCGTCAGCGAGGTGAGCAACGGCACGTCGAGGAGGGCAGG
CTGCGTCTGGCTGCTGCCTCTTCTGGTCTTGCACCTGCTTCTCAAATTT**TGA**TGTGAGTGCC
ACTTCCCCACCCGGGAAAGGCTGCCGCCACCACCACCAACACAACAGCAATGGCAACAC
CGACAGCAACCAATCAGATATATACAAATGAAATTAGAAGAAACACAGCCTCATGGGACAGA
AATTTGAGGGAGGGGAACAAAGAATACTTTGGGGGGAAAAGAGTTTTAAAAAAGAAATTGAA
AATTGCCTTGCAGATATTTAGGTACAATGGAGTTTTCTTTTCCCAAACGGGAAGAACACAGC
ACACCCGGCTTGGACCCACTGCAAGCTGCATCGTGCAACCTCTTTGGTGCCAGTGTGGGCAA
GGGCTCAGCCTCTCTGCCCACAGAGTGCCCCACGTGGAACATTCTGGAGCTGGCCATCCCA
AATTCAATCAGTCCATAGAGACGAACAGAATGAGACCTTCCGGCCCAAGCGTGGCGCTGCGG
GCACTTTGGTAGACTGTGCCACCACGGCGTGTGTTGTGAAACGTGAAATAAAAAGAGCAAAA
AAAAA

52/75

FIGURE 52

MKTIQPKMHNSISWAI FTGLAALCLFQGVFVRSGDATFPKAMDNVTVRQGESATLRCTIDNR
VTRVAWLNRSTILYAGNDKWCLDPRVVLLSNTQTQYSIEIQNVDVYDEGPYTCSVQTDNHPK
TSRVHLIVQVSPKIVEISSDISINEGNNISLTCIATGRPEPTVTWRHISPKAVGFVSEDEYL
EIQGITREQSGDYEC SASNDVAAPVVRVKVTVNYPPISEAKGTGVPVGQKGT LQCEASAV
PSAEFQWYKDDKRLIEGKKGVKVENRPFLSKLIFFNVSEHDYGN YTCVASNKLGH TNASIML
FGPGAVSEVSNGTSRRAGCVWLLPLLVLHLLLKF

Signal peptide:

amino acids 1-28

[illegible]

MKFQGPLACLLLLALCLGSGEAGPLQSGEESTGTNIGEALGHGLGDALSEGVGKAIGKEAGGA
 AGSKVSEALGQGTREAVGTGVRQVPGFGAADALGNRVGEAAHALGNTGHEIGRQAEDVIRHG
 ADAVRGSWQGVPGHSGAWETSGGHGIFGSQGGLGGQGGQGNPGGLGTPWVHGYPGNSAGSFGM
 NPQGAPWGQGGNGGPPNFGTNTQGAVAQPGYGSVRASNQNEGCTNPSPSGSGGGSSNSGGGS
 GSQSGSSGSGSNGDNNNGSSSGSSSGSSSGSSSGSSGGSSGGSSGGSSGNSGGSRGDSGSESSW
 GSSTGSSSGNHGSGGGNGHKPGCEKPGNEARGSGESGIQGFRGQGVSSNMREISKEGNRLL
 GSGDNYRGQGSWSGGGDAVGGVNTVNSETSPGMFNFDTFWKNFKSKLGFINWDAINKDQ
 RSSRIP

55/75

FIGURE 55

AGCCAGGCAGCACATCACAGCGGGAGGAGCTGTCCCAGGTGGCCCAGCTCAGCA**ATGG**CAAT
GGGGGTCCCCAGAGTCATTCTGCTCTGCCTCTTTGGGGCTGCGCTCTGCCTGACAGGGTCCC
AAGCCCTGCAGTGCTACAGCTTTGAGCACACCTACTTTGGCCCCTTTGACCTCAGGGCCATG
AAGCTGCCCAGCATCTCCTGTCCTCATGAGTGCTTTGAGGGCTATCCTGTCTCTGGACACCGG
GTATCGCGCGCCGGTGACCCTGGTGCGGAAGGGCTGCTGGACCGGGCCTCCTGCGGGGCCAGA
CGCAATCGAACCCGGACGCGCTGCCGCCAGACTACTCGGTGGTGCGCGGGCTGCACAACCTGAC
AAATGCAACGCCCCACCTCATGACTCATGACGCCCTCCCCAACCTGAGCCAAGCACCCGACCC
GCCGACGCTCAGCGGGCGCCGAGTGCTACGCCTGTATCGGGGTCCACCAGGATGACTGCGCTA
TCGGCAGGTCCCGACGAGTCCAGTGTCACCAGGACCAGACCGCCTGCTTCCAGGGCAGTGGC
AGAATGACAGTTGGCAATTTCTCAGTCCCTGTGTACATCAGAACCTGCCACCGGCCCTCCTG
CACCACCGAGGGCACCACCAGCCCCCTGGACAGCCATCGACCTCCAGGGCTCCTGCTGTGAGG
GGTACCTCTGCAACAGGAAATCCATGACCCAGCCCTTACCAGTGCTTCAGCCACCACCCCT
CCCCGAGCACTACAGGTCCTGGCCCTGCTCCTCCCAGTCCTCCTGCTGGTGGGGCTCTCAGC
ATAGACCGCCCCCTCCAGGATGCTGGGGACAGGGCTCACACACCTCATTCCTTGCTGCTTCAGC
CCCTATCACATAGCTCACTGGAAAATGATGTTAAAGTAAGAATTGCAAAA

56/75

FIGURE 56

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA86576
><subunit 1 of 1, 251 aa, 1 stop
><MW: 26935, pI: 7.42, NX(S/T): 2
MAMGVPRVILLCLFGAALCLTGSQALQCYSFEHTYFGPFDLRAMKLPSISCPHECFEAILSL
DTGYRAPVTLVRKGCWTGPPAGQTQSNPDALPPDYSVVRGCTTDKCNAHLMTHDALPNLSQA
PDPPTLSGAECYACIGVHQDDCAIGRSRRVQCHQDQTACFQGSGRMTVGNFSPVYIRTCHR
PSCTTEGTTSPWTAIDLQGSCEGYLCNRKSMTQPFTSASATTPPRALQVLALLLPVLLLVG
LSA
```

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 233-251

N-glycosylation sites.

amino acids 120-124, 174-178

N-myristoylation sites.

amino acids 15-21, 84-90

57/75

FIGURE 57

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTCGCGCAGCCTCG
GCACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCTGACTCCGTCCCGGCCAGGGAGGGC
CATGATTTCCCTCCCGGGGCCCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGA
GTGCCCTCGCGCCCCCTCGCGGGCCCAGCTGCAACTGCACTTGCCCGCCAACCGGTTGCAG
GCGGTGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGACGCGGGAGGTGTCTTC
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC
AGGTGTTGTCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT
TAGAACTCAATGTACTGGTTCCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT
GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG
GGTCTTTAAGCCTCACCAACCTTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCAC
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC
AGTGGTTGCTGGAGCTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC
TCTTGTACCACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC
ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT
TTCCTCTGTCACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT
TGACCCCCACGCCCAGTCTCTCCAGCCAGGCCCTGCCCTCACCAAGACTGCCACGACAGAT
GGGGCCCACCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTCTTCCTCTGGCTTGAGCCG
CATGGGTGCTGTGCCTGTGATGGTGCCTGCCCAGAGTCAAGCTGGCTCTCTGGTAT**TGA**TGAC
CCCACCACTCATTGGCTAAAGGATTTGGGGTCTCTCCTTCCTATAAGGGTCACCTCTAGCAC
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCACCTCTC
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG
CTGAAATTAGCTACTACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCAGGC
CCCCTTGATCTGTACCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT
ATTGATATAACCTGTGAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC
TTATTAAACTAACATGAAATATGTGTTGTTTTTCATTTGCAAATTTAAATAAAGATACATAA
TGTTTGTATGAAAAA

58/75

FIGURE 58

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLP AWYTLHGEVSS
SQPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMP SRNLSLRLEGLQEKD SGPY
SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLS CQSPRSKPAVQYQ
WDRQLPSFQTF FAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQC NVTLEVSTGPGAA
VVAGAVVGTTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTL PWP KSSDTISKNGTL
SSVTSARALRPPHGP RP GALTPTPSLSSQALPSPRLPTTDGAHPQPISP I PGGVSSSGLSR
MGAVPVMVPAQSQA GSLV

Signal peptide:

amino acids 1-29

Transmembrane domain:

amino acids 245-267

N-glycosylation site.

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

N-myristoylation site.

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,
262-268, 308-314, 363-369, 364-370

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 164-175

FIGURE 59

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAG
 GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC
 CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT
 CATTTTGATTTTGCTGTTTATTTTTTTTTTCTTTTTCTTTTTCCCACCACATTGTATTTTAT
 TTCCGTACTTCAGAAATGGGCCTACAGACCACAAAGTGGCCCAGCCATGGGGCTTTTTTCCT
 GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAACTCCTGGCCTGCC
 CTAGTGTGTGCCGCTGCGACAGGAACCTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG
 CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAATTAATAATGC
 TGGATTTTCCTGCAGAACTGCACAATGTACAGTCGGTGACACGGTCTACCTGTATGGCAACC
 AACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC
 AATATTTCAGACCATTTACGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGCTGCACCT
 GGATGACAACCTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGAGGCTATTAGCC
 TCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCCTGTGGAC
 TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTCTATATCCGACATGGCCTTCCAGAA
 TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG
 AGGGCACCTTCAGCCATCTCACCAAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC
 CACCCTCCTCCCGATCTCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT
 AAACCACATTCCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA
 ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC
 ACTGCTCGGAATAACCCTTGTTTTGTGACTGCAGTATTAAATGGGTACAGAATGGCTCAA
 ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCTGAACAAGTCCGGG
 GGATGGCCGTCAGGGAATTAAATATGAATCTTTTGTCTGTCCCACCACGACCCCCGGCCTG
 CCTCTCTTCACCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTAT
 TCCAAACCCTAGCAGAAGCTACACGCCTCCAACCTCCTACCACATCGAAACTTCCACGATT
 CTGACTGGGATGGCAGAGAAAGAGTGACCCACCTATTTCTGAACGGATCCAGCTCTCTATC
 CATTTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTCACCGTGATGGCATA
 CAAACTCACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTTCAGGAGCGCATAG
 TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT
 TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTTCAGAGGC
 CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA
 CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTT
 GTGCTGGTGGTCTTGCTCAGCGTCTTTTGCTGGCATATGCACAAAAGGGGCGCTACACCTC
 CCAGAAGTGGAATACAACCGGGGGCGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGA
 AGGACAACTCCATCCTGGAGATGACAGAAACCAGTTTTTCAGATCGTCTCCTTAAATAACGAT
 CAACTCCTTAAAGGAGATTTTCAGACTGCAGCCCATTTACACCCCAAATGGGGGCATTAATTA
 CACAGACTGCCATATCCCCAACACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC
 ACTGCCATACGTGACAGCCAGAGGGCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA
 CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT
 TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAGTG
 CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTTGTAACCTTTTGCTTTTTAAATCTT

FIGURE 60

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP
 EGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTI
 SRAALAQLLKLEELHLDDNSISTVGVEDGAFREAI SLKLLFLSKNHLSSVPVGLPVDLQELR
 VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD
 LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN
 PWFCDCSIKWVTEWLKYIPSSLNVRGFMCGPEQVRGMVRELNMNLLSCPTTTPGLPLFTP
 APSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPPISERIQLSIHFVND
 TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL
 DAFNYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSMSGSPFLLAGLIGGAVIFVLVVL
 LSVFCWHMHKKGRYTSQKWKYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG
 DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

Signal peptide:

amino acids 1-42

Transmembrane domain:

amino acids 542-561

N-glycosylation site.

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

Casein kinase II phosphorylation site.

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

Tyrosine kinase phosphorylation site.

amino acids 319-328

N-myristoylation site.

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,
 522-528, 545-551, 633-639

Amidation site.

amino acids 581-585

Leucine zipper pattern.

amino acids 164-186

Phospholipase A2 aspartic acid active site.

amino acids 39-50

62/75

FIGURE 62

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906
><subunit 1 of 1, 772 aa, 1 stop
><MW: 87002, pI: 4.64, NX(S/T): 8
MNCYLLLRFMLGIPLLWPCLGATENSQTKKVKQPVRSHLRVKRGVWNQFFVPEEMNTTSHH
IGQLRSDLDNGNNSFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIAT
GRAVEPESEFVIKVSNDINDNEPKFLDEPYEAIVPEMSPEGTLVIQVTASDADDPSSGNNARL
LYSLLQGQPYFSVEPTTGIVIRISSKMDRELQDEYVWIIQAKDMIGQPGALSGTTSVLIKLSD
VNDNKPIFKESLYRLTVSESAPTGTSGTIGTIMAYDNDIGENAEMDYSIEEDDSQTFDIITNHE
TQEGIVILKKKVD FEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLL
PYYVFEVFEETPQGSFVGVSATDPDNRKSPIRYSITRSKVFENINDNGTITTSNSLDREISA
WYNLSITATEKYNIEQISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTISAVDR
DESIEEHFHYFNLSVEDTNNSSFTIIDNQDNTAVILTNRTGFNLQEEPVFYISILIADNGIP
SLTSTNTLTIHVCDGDSGSTQTCQYQELVLSMGFKTEVIIAILICIMIIIFGFIFLTLGLKQ
RRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTSAEIRSLY
RQSLQVGPDSAIFRKFILEKLEEANTDPCAPPFDSLQTYAFEGTGS LAGSLSSLES AVSDQD
ESYDYLNELGPRFKRLACMFGSAVQSNN
```

Important features:

Signal peptide:

amino acids 1-21

Transmembrane domain:

amino acids 597-617

N-glycosylation sites.

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518, 516-519 and 534-537

Cadherins extracellular repeated domain signature.

amino acids 136-146 and 244-254

FIGURE 63

CTTCAGAACAGGTTCTCCTTCCCCAGTCACCAGTTGCTCGAGTTAGAATTGTCTGCA**ATGGC**
CGCCCTGCAGAAATCTGTGAGCTCTTTCCTTATGGGGACCCTGGCCACCAGCTGCCTCCTTC
TCTTGGCCCTCTTGGTACAGGGAGGAGCAGCTGCGCCCATCAGCTCCCAGTGCAGGCTTGAC
AAGTCCAACCTCCAGCAGCCCTATATCACCAACCGCACCTTCATGCTGGCTAAGGAGGCTAG
CTTGGCTGATAACAACACAGACGTTTCGTCTCATTGGGGAGAACTGTTCCACGGAGTCAGTA
TGAGTGAGCGCTGCTATCTGATGAAGCAGGTGCTGAACCTCACCCCTGAAGAAGTGCTGTTC
CCTCAATCTGATAGGTTCCAGCCTTATATGCAGGAGGTGGTGGCCCTTCCTGGCCAGGCTCAG
CAACAGGCTAAGCACATGTCATATTGAAGGTGATGACCTGCATATCCAGAGGAATGTGCAAA
AGCTGAAGGACACAGTGAAAAAGCTTGGAGAGAGTGAGAGATCAAAGCAATTGGAGAAGTG
GATTTGCTGTTTATGTCTCTGAGAAATGCCTGCATT**TGA**CCAGAGCAAAGCTGAAAAATGAA
TAACTAACCCCTTTCCCTGCTAGAAATAACAATTAGATGCCCCAAAGCGATTTTTTTTAAAC
CAAAAGGAAGATGGGAAGCCAACTCCATCATGATGGGTGGATTCCAAATGAACCCCTGCGT
TAGTTACAAAGGAAACCAATGCCACTTTTGTTTATAAGACCAGAAGGTAGACTTTCTAAGCA
TAGATATTTATTGATAACATTTTCATTGTAAGTGGTGTCTATACACAGAAAACAATTTATTT
TTTAAATAATTGTCTTTTCCATAAAAAAGATTACTTTCCATTCCTTTAGGGGAAAAAACCC
CTAAATAGCTTCATGTTTCCATAATCAGTACTTTATATTTATAAATGTATTTATTATTATTA
TAAGACTGCATTTTATTTATATCATTTTATTAATATGGATTTATTTATAGAAACATCATTCG
ATATTGCTACTTGAGTGTAAGGCTAATATTGATATTTATGACAATAATTATAGAGCTATAAC
ATGTTTATTTGACCTCAATAAACACTTGGATATCCC

64/75

FIGURE 64

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125185

><subunit 1 of 1, 179 aa, 1 stop

><MW: 20011, pI: 8.10, NX(S/T): 3

MAALQKSVSSFLMGTLATSCLLLLLALLVQGGAAAPISSHCRLDKSNFQQPYITNRTFMLAKE
ASLADNNTDVRLIGEKLFFHGVSMSERCYLMKQVLNFTLEEVLFPPQSDRFQPYMQEVVPFLAR
LSNRLSTCHIEGDDLHIQRNVQKLKDTVKKLGESGEIKAIGELDLLFMSLRNACI

Important features of the protein:

Signal peptide:

amino acids 1-33

N-glycosylation sites.

amino acids 54-58, 68-72, 97-101

N-myristoylation sites.

amino acids 14-20, 82-88

Prokaryotic membrane lipoprotein lipid attachment site.

amino acids 10-21

65/75

FIGURE 65

GCCCTAACCTTCCCAGGGCTCAGCTCTTTGGAGCTGCCCATTCCTCCGGCTGCGAGAAAGGA
CGCGCGCCCTGCGTCGGGCGAAGAAAAGAAGCAAACTTGTCTGGGAGGGTTTCGTCATCAAC
CTCCTTCCCGCAAACCTAAACCTCCTGCCGGGGCCATCCCTAGACAGAGGAAAGTTCCTGCA
GAGCCGACCAGCCCTAGTGGATCTGGGGCAGGCAGCGGGCTGGCTGTGGAATTAGATCTGT
TTTGAACCCAGTGGAGCGCATCGCTGGGGCTCGGAAGTCACCGTCCGCGGGCACCGGGTTGG
CGCTGCCCCGAGTGGAAACCGACAGTTTGCGAGCCTCGGCTGCAAGTGGCCTCTCCTCCCCCG
GTTGTTGTTCACTGTCGGGTGAGGGCTGCGAGTGTGGCAAGTTGCAAGAGAGCCTCAGAGG
TCCGAAGAGCGCTGCGCTCCTACTCGCGTTCTGCTTCTTCTCTCGGTTCCCTACTGTGA
AATCGCAGCGACATTTACAAAGGCCCTCCGGGTCTTACCGAGACCGATCCGCAGCGTTTGGCC
CGGTCTGCTTATTGCATCGGGAGCCCCGAGCACCGGGCGAAATGCGGAGGTTCCCGAAGGC
CGACCTGGCCGCTGCAGGAGTTATGTTACTTTGCCACTTCTTACGGACAGTTTTCAGTTTCG
CCGATGGGAACCCGGAGACCAATCCTTGATTGGCAGTATGGAGTTACTCAGGCCCTTCCCT
CACACAGAGGAGGAGGTGGAAGTTGATTACACGCGTACAGCCACAGGTGGAAAAGAACTT
GGACTTTCTCAAGGCGGTAGACACGAACCGAGCAAGCGTCGGCCAAGACTCTCCTGAGCCCA
GAAGCTTCACAGACCTGCTGCTGGATGATGGGCAGGACAATAACACTCAGATCGAGGAGGAT
ACAGACCACAATTACTATATATCTCGAATATATGGTCCATCTGATTCTGCCAGCCGGGATTT
ATGGGTGAACATAGACCAAATGGAAAAAGATAAAGTGAAGATTCATGGAATATTGTCCAATA
CTCATCGGCAAGCTGCAAGAGTGAATCTGTCCTTCGATTTTCCATTTTATGGCCACTTCCTA
CGTGAAATCACTGTGGCAACCGGGGGTTTCATATACACTGGAGAAGTCGTACATCGAATGCT
AACAGCCACACAGTACATAGCACCTTTAATGGCAAATTTTCGATCCCAGTGTATCCAGAAATT
CAACTGTCAGATATTTTGATAATGGCACAGCACTTGTGGTCCAGTGGGACCATGTACATCTC
CAGGATAATTATAACCTGGGAAGCTTCACATTCCAGGCAACCCTGCTCATGGATGGACGAAT
CATCTTTGGATACAAAGAAATTCCTGTCTTGGTCACACAGATAAGTTCAACCAATCATCCAG
TGAAAGTCGGACTGTCCGATGCATTTGTCTGTGCCACAGGATCCAACAAATTCCCAATGTT
CGAAGAAGAACAATTTATGAATACCACCGAGTAGAGCTACAAATGTCAAAAATTACCAACAT
TTCGGCTGTGGAGATGACCCCATACCCACATGCCTCCAGTTTAAACAGATGTGGCCCCCTGTG
TATCTTCTCAGATTGGCTTCAACTGCAGTTGGTGTAGTAACTTCAAAGATGTTCCAGTGGA
TTTGATCGTCATCGGCAGGACTGGGTGGACAGTGGATGCCCTGAAGAGTCAAAAGAGAAGAT
GTGTGAGAATACAGAACCAGTGGAAACTTCTTCTCGAACCACCACAACCGTAGGAGCGACAA
CCACCCAGTTCAGGGTCCTAACTACCACCAGAAGAGCAGTGACTTCTCAGTTTCCCACCAGC
CTCCCTACAGAAGATGATACCAAGATAGCACTACATCTAAAAGATAATGGAGCTTCTACAGA
TGACAGTGCAGCTGAGAAGAAAGGGGGAACCTCCACGCTGGCCTCATCATTGGAATCCTCA
TCCTGGTCCCTCATTGTAGCCACAGCCATTCTTGTGACAGTCTATATGTATCACCACCCAACA
TCAGCAGCCAGCATCTTCTTTATTGAGAGACGCCCAAGCAGATGGCCTGCGATGAAGTTTAG
AAGAGGCTCTGGACATCCTGCCTATGCTGAAGTTGAACCAGTTGGAGAGAAAGAAGGCTTTA
TTGTATCAGAGCAGTGCTTAAATTTCTAGGACAGAACAACACCAGTACTGGTTTACAGGTGT
TAAGACTAAAATTTTGCCTATACCTTTAAGACAAACAAACACACACACAAACAAGCTC
TAAGCTGCTGTAGCCTGAAGAAGACAAGATTTCTGGACAAGCTCAGCCCAGGAAACAAAGGG
TAAACAAAAAACTAAACTTATACAAGATACCATTTTACTGAACATAGAATTCCCTAGTGG
AATGTCATCTATAGTTCCTCGGAACATCTCCCGTGGACTTATCTGAAGTATGACAAGATTA
TAATGCTTTTGGCTTAGGTGCAGGGTTGCAAAGGGATCAGAAAAAAAATCATAATAAAGC
TTTAGTTCATGAGGG

FIGURE 66

MARFPKADLAAAGVMLLCHFFTDQFQFADGKPGDQILDWQYGVTAQAFPHTEEEVEVD SHAYS
HRWKRNLDFLKAVDTNRASVGQDSPEPRSFTDLLLDDGQDNNTQIEEDTDHNYIISRIYGPS
DSASRDLWVNIDQMEKDKVKI HGILSNTHRQAARVNLSFDFPFYGHFLREITVATGGFIYTG
EVVHRMLTATQYIAPLMANFDPSVSRNSTVRYFDNGTALVVQWDHVHLQDNYNLGSFTFQAT
LLMDGRIIFGYKEIPVLVTQISSTNHPVKVGLSDAFVVVHRIQQIPNVRRRTIYEYHRVELQ
MSKITNISAVEMTPLPTCLQFNRCGPCVSSQIGFNC SWCSKLQRCSSGFDRHRQDWVDSGCP
EESKEKMCENTEPVETSSRTTTTVGATTTQFRVLTTTRRAVTSQFPTSLPTEDDTKIALHLK
DNGASTDDSAAEKKGGTLHAGLIIGILILVLIVATAILVTVYMYHHPTSAASIFFIERRPSR
WPAMKFRRGSGHPAYAEVEPVGEKEGFIVSEQC

Important features of the protein:**Transmembrane domain:**

amino acids 454-478

N-glycosylation sites.

amino acids 103-107, 160-164, 213-217, 221-225, 316-320, 345-349

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 297-301, 492-496, 503-507

N-myristoylation sites.

amino acids 42-48, 100-106, 147-153, 279-285, 397-403, 450-456,
455-461

GACGACCCCTAGCAGGGATGCGACATGATGCTGTTGGTGCAGGGTGCTTGTGCTCGAACCCAGTG
 GCTGGCGGCGGCTGCTCCTCAGCCTGTGCTGCCTGCTACCCCTCCTGCCTCCCGGCTGGACAGA
 GTGTGGACTTCCCCCTGGGCGGCCGTGGACAACATGATGGTCAGAAAAGGGGACACGGCGGTG
 CTTAGGTGTTATTTTGAAGATGGAGCTTCAAAGGGTGCCTGGCTGAACCGGTCAAGTATTAT
 TTTTGGCGGGAGGTGATAAGTGGTCAGTGGATCCTCGAGTTTCAATTTCAACATTGAATAAAA
 GGGACTACAGCCTCCAGATACAGAATGTAGATGTGACAGATGATGGCCCATACACGTGTTCT
 GTTCAGACTCAACATACACCCAGAACAATGCAGGTGCATCTAACTGTGCAAGTTCCTCCTAA
 GATATATGACATCTCAAATGATATGACCGTCAATGAAGGAACCAACGTCACTCTTACTTGTT
 TGGCCACTGGGAAACCAGAGCCTTCCATTTCTTGGCGACACATCTCCCCATCAGCAAAACCA
 TTTGAAAATGGACAATATTTGGACATTTATGGAATTACAAGGGACCAGGCTGGGGAATATGA
 ATGCAGTGCGGAAAATGATGTGTCATTCCCAGATGTGAGGAAAGTAAAAGTTGTTGTCAACT
 TTGCTCCTACTATTTCAGGAAATTAAATCTGGCACCGTGACCCCCGGACGCAGTGGCCTGATA
 AGATGTGAAGGTGCAGGTGTGCCGCCTCCAGCCTTTGAATGGTACAAAGGAGAGAAGAAGCT
 CTTCAATGGCCAACAAGGAATTATTATTCAAATTTTAGCACAAGATCCATTCTCACTGTTA
 CCAACGTGACACAGGAGCACTTCGGCAATTATACTTGTGTGGCTGCCAACAAGCTAGGCACA
 ACCAATGCGAGCCTGCCTCTTAACCCTCCAAGTACAGCCCAGTATGGAATTACCGGGAGCGC
 TGATGTTCTTTTCTCCTGCTGGTACCTTGTGTTGACACTGTCCTCTTTCACCAGCATATTCT
 ACCTGAAGAATGCCATTCTACAATTAAATTCAAAGACCCATAAAAGGCTTTTAAGGATTCTCT
 GAAAGTGCTGATGGCTGGATCCAATCTGGTACAGTTTGTATAAAAGCAGCGTGGGATATAATC
 AGCAGTGCTTACATGGGGATGATCGCCTTCTGTAGAATTGCTCATTATGTAAATACTTTAAT
 TCTACTCTTTTTTTGATTAGCTACATTACCTTGTGAAGCAGTACACATTGTCCTTTTTTTAAG
 ACGTGAAAGCTCTGAAATTACTTTTTAGAGGATATTAATTGTGATTTTCATGTTTGTAAATCTAC
 AACTTTTTCAAAGCATTTCAGTCATGGTCTGCTAGGTTGCAGGCTGTAGTTTACAAAACGAA
 TATTGCAGTGAATATGTGATTCTTTAAGGCTGCAATACAAGCATTTCAGTTCCCTGTTTTCAAT
 AAGAGTCAATCCACATTTACAAAGATGCATTTTTTTCTTTTTTGATAAAAAAGCAAATAATA
 TTGCCTTTCAGATTATTTCTTCAAATATAACACATATCTAGATTTTTCTGCTCGCATGATAT
 TCAGGTTTTCAGGAATGAGCCTTGTAATATAACTGGCTGTGCAGCTCTGCTTCTCTTTCCTGT
 AAGTTCAGCATGGGTGTGCCTTTCATACAATAATATTTTTCTCTTTGTCTCCAATAATATAA
 AATGTTTTTGCTAAATCTTACAATTTGAAAGTAAAAATAAACCAGAGTGATCAAGTTAAACCA
 TACACTATCTCTAAGTAACGAAGGAGCTATTGGACTGTAAAAATCTCTTCCTGCACTGACAA
 TGGGGTTTTGAGAATTTTGCCCCACACTAACTCAGTTCCTTGTGATGAGAGACAATTTAATAAC
 AGTATAGTAAATATACCATATGATTTCTTTAGTTGTAGCTAAATGTTAGATCCACCGTGGGA
 AATCATTCCCTTTAAATGACAGCACAGTCCACTCAAAGGATTGCCTAGCAATACAGCATCT
 TTTCTTTTCACTAGTCCAAGCCAAAAATTTTAAGATGATTTGTGAGAAAGGGCACAAAGTCC
 TATCACCTAATATTACAAGAGTTGGTAAGCGCTCATCATTATTTTATTTTGTGGCAGCTAA
 GTTAGTATGACAGAGGCAGTGCTCCTGTGGACAGGAGCATTTTGCATATTTTCCATCTGAAA
 GTATCACTCAGTTGATAGTCTGGAATGCATGTTATATATTTTAAAACTTCCAAAATATATTA
 TAACAAACATTTCTATATCGGTATGTAGCAGACCAATCTCTAAATAGCTAATTCTTCAATAA
 AATCTTTCTATATAGCCATTTTCAGTGCAAAACAAGTAAATCAAAAAAGACCATCTTTATTT
 TTCTTTACATGATATATGTAAGATGCGATCAAATAAAGACAAACACCAGTGATGAGAAATAT
 CTTAAGATAAGTAATTATCAAATTATTGTGAATGTTAAATTATTTCTACTATAAAGAAGCAA
 AACTACATTTTTTGAAGGAAAATGCTGTTACTCTAACATTAATTTACAGGAATAGTTTGTATGG
 TTTCACTCTTTACTAAAGAAAGGCCATCACCTTGAAAGCCATTTTACAGGTTTGATGAAGTT
 ACCAATTTTCAGTACACCTAAATTTCTACAAATAGTCCCCTTTTACAAGTTGTAACAACAAAG
 ACCCTATAATAAAAATTAGATACAAGAAATTTTGCAGTGGTTATACATATTTGAGATATCTAG
 TATGTTGCCCTAGCAGGGATGGCTTAAAAACTGTGATTTTTTTTTCTTCAAGTAAAACCTTAGT
 CCCAAAGTACATCATAAATCAATTTTAATTAGAAAAATGAATCTTAAATGAGGGGACATAAG
 TATACTCTTTCCACAAAATGGCAATAATAAGGCATAAAGCTAGTAAATCTACTAACTGTAAT
 AAATGTATGACATTATTTTGTATTGATACATTAAAAAAGAGTTTTTTAGAACAAATATGGCATT
 TAACTTTATTATTTATTTGCTTTTAAAGAAATATTCTTTGTGGAATTGTTGAATAAACTATAA
 AATATTATTTTGTATTGCAGCTTTAAAGTGGCACACTCCATAATAATCTACTTACTAGAAAT

FIGURE 67B

AGTGGTGCTACCACAAAAAATGTTAACCATCAGTACCATTGTTTGGGAGAAAGAAACAGATC
AAGAATGCATATTATTTCAGTGACCGCTTTCCTAGAGTTAAAATACCTCCTCTTTGTAAGGTT
TGTAGGTAAATTGAGGTATAAACTATGGATGAACCAAATAATTAGTTCAAAGTGTTGTCATG
ATTCCAAATTTGTGGAGTCTGGTGTTTTTACCATAGAATGTGACAGAAGTACAGTCATAGCT
CAGTAGCTATATGTATTTGCCTTTATGTTAGAAGAGACTTTCCTTGAGTGACATTTTTTAAATA
GAGGAGGTATTCACTATGTTTTTCTGTATCACAGCAGCATTCCTAGTCCTTAGGCCCTCGGA
CAGAGTGAAATCATGAGTATTTATGAGTTCAATATTGTCAAATAAGGCTACAGTATTTGCTT
TTTTGTGTGAATGTATTGCATATAATGTTCAAGTAGATGATTTTACATTTATGGACATATAA
AATGTCTGATTACCCCATTTTATCAGTCCTGACTGTACAAGATTGTTGCAATTTTCAGAATAG
CAGTTTTATAAATTGATTTATCTTTTAATCTATAACAATTTGTGTTAGCTGTTTCATTTTCAGG
ANTATATTTTCTACAAGTTCCACTTGTGGGACTCCTTTTGTGCCCCCTATTTTTTTTTTAAAG
AAGGAAGAAAGAAAAATAAGTAGCAGTTTAAAAATGAGAATGGAGAGAAAAGAAAAAGAATG
AAAAGGAAAGGCAGTAAAGAGGGAAAAAAAAGGAAGGATGGAAGGAATGAAGGAAGGAAGGG
AGGAAGGGGAGAAGGTAGGAAGAAAGAAAGGATGAGAGGGAAGGAAGAATCAGAGTATTAGG
GTAGTTAACTTACACATTTGCATTCTTAGTTTAACTGCAAGTGGTGTAAGTATGTTTTTCAA
TGATCGCATTTGAAACATAAGTCCTATTATACCATTAAAGTTCCATTATGCAGCAATTATAT
AATAAAAAGTACTGCCCAAGTTATAGTAATGTGGGTGTTTTTGAGACACTAAAAGATTTGAG
AGGGAGAATTTCAAACCTTAAAGCCACTTTTGGGGGGTTTATACTTAACTGAAAAATTAATG
CTTCATCATAACATTTAAGCTATATCTAGAAAGTAGACTGGAGAACTGAGAAAATTACCCAG
GTAATTCAGGGAAAAAAAATATATATATATATAAATACCCCTACATTTGAAGTCAGAAA
ACTCTGAAAACTGAATTATCAAAGTCAATCATCTATAATGATCAAATTTACTGAACAATTG
TTAATTTATCCATTGTGCTTAGCTTTGTGACACAGCCAAAAGTTACCTATTTAATCTTTTCA
ATAAAAATTGTTTTTTGAAATCCAGAAATGATTTAAAAAGAGGTCAGGTTTTTAACTATTTA
TTGAAGTATGTGGATGTACAGTATTTCAATAGATATGAATATGAATAAATGGTATGCCTTAA
GATTCTTTGAATATGTATTTACTTTAAAGACTGGAAAAAGCTCTTCCTGTCTTTTAGTAAAA
CATCCATATTTCATAACCTGATGTAAAATATGTTGTACTGTTTCCAATAGGTGAATATAAAC
TCAGTTTATCAATTAAAAA

FIGURE 68

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA92259
><subunit 1 of 1, 354 aa, 1 stop
><MW: 38719, pI: 6.12, NX(S/T): 6
MDMMLLVQGACCSNQWLAHAVLLSLCCLLPAGQSVDFPWAAVDNMMVRKGD TAVLR CYL
EDGASKGAWLNRSSIIIFAGGDKWSVDPRVSISTLNKR DYSLQIQNV DVTDDGPYTCSVQTQH
TPRTMQVHLTVQVPPKIYDISNDMTVNEGTVTLTCLATGKPEPSISWRHISPSAKPFENGQ
YLDIYGITRDQAGEYEC SAENDVSFPDVRKVKVVVNFAPTIQEIKSGTVTPGRSGLIRCEGA
GVPPPAFEWYKGEKKLFNGQQGIIIIQN FSTRSILTVTNVTQEHFGNYTCVAANKLGTTNASL
PLNPPSTAQYGITGSADVLFSCWYLVLTLS SFTSIFYLKNAILQ
```

Important features of the protein:**Signal peptide:**

amino acids 1-33

Transmembrane domain:

amino acids 322-343

N-glycosylation sites.

amino acids 73-77, 155-159, 275-279, 286-290, 294-298, 307-311

Tyrosine kinase phosphorylation site.

amino acids 180-188

N-myristoylation sites.amino acids 9-15, 65-71, 69-75, 153-159, 241-247, 293-299,
304-310, 321-327**Myelin P0 protein.**

amino acids 94-123

70/75

FIGURE 69

ATAGTAGAAGAATGTCTCTGAAATTACTGGATGAGTTTCAGTCATACTTTCACATGGGCACA
ATTTACATTCAAGCTCCTTATCCTAGGCTAATTTTATATTATGTTAAATCACTTGTTTTTG
TTCTCACGGCTTCCTGCCTGCTATAGGCATAATTACGAGGAAGCAGAACTTCTCCAGAAGCA
AGCGCACATGCGTTCCAAAATAAGAGCAAATTCGCTCTAAACACAGGAAAAGACCTGAAGCT
TTAATTAAGGGGTACATCCAACCCAGAGCGCTTTTGTGGGCACTGATTGCTCCAGCTTCT
GCGTCACTGCGCGAGGGAAGAGGGAAGAGGATCCAGGCGTTAGAC**ATG**TATAGACACAAAAA
CAGCTGGAGATTGGGCTTAAATAACCCACCAAGCTCCAAAGAAGAGACCCAAGTCCCCAAAA
CATTGATTTTCAGGGCTGCCAGGAAGGAAGAGCAGCAGCAGGGTGGGAGAGAAGCTCCAGTCA
GCCCACAAGATGCCATTGTCCCCCGGCCTCCTGCTGCTGCTGCTCTCCGGGGCCACGGCCAC
CGCTGCCCTGCCCCTGGAGGGTGGCCCCACCGGCCGAGACAGCGAGCATATGCAGGAAGCGG
CAGGAATAAGGAAAAGCAGCCTCCTGACTTTCCTCGCTTGGTGGTTTGAGTGGACCTCCCAG
GCCAGTGCCGGGCCCCTCATAGGAGAGGAAGCTCGGGAGGTGGCCAGGCGGCAGGAAGGCGC
ACCCCCCAGCAATCCGCGCGCCGGGACAGAATGCCCTGCAGGAACTTCTTCTGGAAGACCT
TCTCCTCCTGCAAA**TAG**

71/75

FIGURE 70

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA44175

><subunit 1 of 1, 155 aa, 1 stop

><MW: 17194, pI: 10.44, NX(S/T): 0

MYRHKNSWRLGLKYPPSSKEETQVPKTLISGLPGRKSSSRVGEKLQSAHKMPLSPGLLLLLL
SGATATAALPLEGGPTGRDSEHMQEAAAGIRKSSLLTFLAWWFEWTSQASAGPLIGEEAREVA
RRQEGAPPQQSARRDRMPCRNFFWKTFSSCK

Important features of the protein:

Transmembrane domain:

amino acids 51-69

cAMP- and cGMP-dependent protein kinase phosphorylation sites.

amino acids 35-39, 92-96

N-myristoylation sites.

amino acids 64-70, 75-81, 90-96

Amidation site.

amino acids 33-37

72/75

FIGURE 71

GTCGTGTGCTTGGAGGAAGCCGCGGAACCCCCAGCGTCCGTCC**ATG**GCGTGGAGCCTTGGGA
GCTGGCTGGGTGGCTGCCTGCTGGTGTGAGCATTGGGAATGGTACCACCTCCCGAAAATGTC
AGAATGAATTCTGTTAATTTCAAGAACATTCTACAGTGGGAGTCACCTGCTTTTGCCAAAGG
GAACCTGACTTTTACAGCTCAGTACCTAAGTTATAGGATATTCCAAGATAAATGCATGAATA
CTACCTTGACGGAATGTGATTTCTCAAGTCTTTCCAAGTATGGTGACCACACCTTGAGAGTC
AGGGCTGAATTTGCAGATGAGCATTGAGACTGGGTAAACATCACCTTCTGTCCTGTGGATGA
CACCATTATTGGACCCCCTGGAATGCAAGTAGAAGTACTTGCTGATTCTTTACATATGCGTT
TCTTAGCCCCCTAAAATTGAGAATGAATACGAAACTTGGACTATGAAGAATGTGTATAACTCA
TGGACTTATAATGTGCAATACTGGAAAAACGGTACTGATGAAAAGTTTCAAATTACTCCCCA
GTATGACTTTTGAGGTCCCTCAGAAACCTGGAGCCATGGACAACCTATTGTGTTCAAGTTCGAG
GGTTTCTTCCTGATCGGAACAAAGCTGGGGAATGGAGTGAGCCTGTCTGTGAGCAAACAACC
CATGACGAAACGGTCCCCCTCCTGGATGGTGGCCGTCATCCTCATGGCCTCGGTCTTCATGGT
CTGCCTGGCACTCCTCGGCTGCTTCTCCTTGCTGTGGTGCCTTTACAAGAAGACAAAGTACG
CCTTCTCCCCCTAGGAATTCTCTTCCACAGCACCTGAAAGAGTTTTTGGGCCATCCTCATCAT
AACACACTTCTGTTTTTCTCCTTTCCATTGTGCGGATGAGAATGATGTTTTTGACAAGCTAAG
TGTCATTGCAGAAGACTCTGAGAGCGGCAAGCAGAATCCTGGTGACAGCTGCAGCCTCGGGA
CCCCGCCTGGGCAGGGGCCCCAAAGC**TAG**GCTCTGAGAAGGAAACACACTCGGCTGGGCACA
GTGACGTACTCCATCTCACATCTGCCTCAGTGAGGGATCAGGGCAGCAAACAAGGGCCAAGA
CCATCTGAGCCAGCCCCACATCTAGAACTCCAGACCTGGACTTAGCCACCAGAGAGCTACAT
TTTAAAGGCTGTCTTGGCAAAAATACTCCATTTGGGAACTCACTGCCTTATAAAGGCTTTCA
TGATGTTTTTCAGAAGTTGGCCACTGAGAGTGTAATTTTCAGCCTTTTATATCACTAAAATAA
GATCATGTTTTAATTGTGAGAAACAGGGCCGAGCACAGTGGCTCACGCCTGTAATACCAGCA
CCTTAGAGGTCGAGGCAGGCGGATCACTTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAATA
TGGTGAAACCCAGTCTCTACTAAAAATACAAAAATTAGCTAGGCATGATGGCGCATGCCTAT
AATCCCAGCTACTCGAGTGCCTGAGGCAGGAGAATTGCATGAACCCGGGAGGAGGAGGAGGA
GGTTGCAGTGAGCCGAGATAGCGGCACTGCACTCCAGCCTGGGTGACAAAGTGAGACTCCAT
CTCAAAAAAAAAAAAAAAAAAATTGTGAGAAACAGAAATACTTAAATGAGGAATAAGAATGG
AGATGTTACATCTGGTAGATGTAACATTCTACCAGATTATGGATGGACTGATCTGAAAATCG
ACCTCAACTCAAGGGTGGTCAGCTCAATGCTACACAGAGCACGGACTTTTGGATTCTTTGCA
GTACTTTGAATTTATTTTTCTACCTATATATGTTTTATATGCTGCTGGTGTCCATTAAAGT
TTTACTCTGTGTTGC

73/75

FIGURE 72

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA83551
><subunit 1 of 1, 325 aa, 1 stop
><MW: 37011, pI: 5.09, NX(S/T): 4
MAWSLGSWLGGCLLVSALGMVPPPENVRMNSVNFKNILQWESPAFAKGNLTFTAQYLSYRIF
QDKCMNTTLTECDFSSLSKYGDHTLRVRAEFADHSDWVNITFCPVDDTIIGPPGMQVEVLA
DSLHMRFLAPKIENEYETWTMKNVYNSWTYNVQYWKNGTDEKFQITPQYDFEVLRLNLEPWTT
YCVQVRGFLPDRNKAGEWSEPVCEQTTTHDETVPSPWMAVILMASVFMVCLALLGCFSLWLVCV
YKKTKYAFSPRNSLPQHLKEFLGHPHNTLLFFSFPLSDENDVFDKLSVIAEDSESGKQNPQ
DSCSLGTTPPGQGPQS
```

Important features of the protein:

Signal peptide:

amino acids 1-19

Transmembrane domain:

amino acids 222-245

N-glycosylation sites.

amino acids 49-53, 68-72, 102-106, 161-165

N-myristoylation sites.

amino acids 6-12, 316-322

FIGURE 73

CGAGCGCCAACCCGCTAGCGCCTGAATCCGGCGTGCTGCCCGCTCGCCGCCCGCC**ATGGCCC**
GCGCAGCCCCGCTGCTCGCCGCGTTGACCGCGCTCCTCGCCGCCCGCGCTGCTGGCGGAGAT
GCCCCGCCGGGCAAAATCGCGGTGGTTGGGGCTGGGATTGGGGGCTCTGCTGTGGCCCATTT
TCTCCAGCAGCACTTTGGACCTCGGGTGAGATCGACGTGTACGAGAAGGGAACCGTGGGTG
GCCGCTTGGCCACCATCTCAGTCAACAAGCAGCACTATGAGAGCGGGGCTGCCTCCTTCCAC
TCCCTGAGCCTGCACATGCAGGACTTCGTCAAGCTGCTGGGGCTGAGGCACCGGCGCGAGGT
GGTGGGCAGGAGCGCCATCTTCGGCGGGGAGCACTTCATGCTGGAGGAGACTGACTGGTACC
TGCTGAACCTCTTCCGCCTCTGGTGGCACTATGGCATCAGCTTCCTGAGGCTGCAGATGTGG
GTGGAGGAGGTGATGGAGAAGTTCATGAGGATCTATAAGTACCAGGCCACGGCTATGCCTT
CTCGGGTGTGGAGGAGCTGCTCTACTCACTGGGGGAGTCCACCTTTGTAAACATGACCCAGC
ACTCTGTGGCTGAGTCCCTGCTGCAGGTGGGCGTCACGCAGCGCTTTATTGATGATGTGCTT
TCTGCTGTCTGCGGGCCAGCTATGGCCAGTCAGCAGCGATGCCCGCCTTTGCAGGAGCCAT
GTCACTAGCCGGGGCCCAAGGCAGCCTGTGGTCTGTGGAAGGAGGCAATAAGCTGGTTTGT
CCGGTTTGTGAAGCTCACCAAGGCCAATGTGATCCATGCCACAGTGACCTCTGTGACCCTG
CACAGCACAGAGGGGAAAGCCCTGTACCAGGTGGCGTATGAGAATGAGGTAGGCAACAGCTC
TGACTTCTATGACATCGTGGTCATCGCCACCCCCCTGCACCTGGACAACAGCAGCAGCAACT
TAACCTTTGCAGGCTTCCACCCGCCCATTTGATGACGTGCAGGGCTCTTTCCAGCCCACCGTC
GTCTCCTTGGTCCACGGCTACCTCAACTCGTCCTACTTCGGTTTCCAGACCCTAAGCTTTT
CCCCTTTGCCAACATCCTTACCACAGATTTCCCCAGCTTCTTCTGCACTCTGGACAACATCT
GCCCTGTCAACATCTCTGCCAGCTTCCGGCGAAAGCAGCCCCAGGAGGCAGCTGTTTGGCGA
GTCCAGTCCCCCAAGCCCCCTCTTTCGGACCCAGCTAAAGACCCTGTTCCGTTCCCTATTACTC
AGTGCAGACAGCTGAGTGGCAGGCCCATCCCCTCTATGGCTCCCGCCCCACGCTCCCGAGGT
TTGCACTCCATGACCAGCTCTTCTACCTCAATGCCCTGGAGTGGGCGGCCAGCTCCGTGGAG
GTGATGGCCGTGGCTGCCAAGAATGTGGCCTTGCTGGCTTACAACCGCTGGTACCAGGACCT
AGACAAGATTGATCAAAAAGATTTGATGCACAAGGTCAAGACTGAACGTGTGAGGGCTCTAGG
GAGAGCCTGGGAACTTTTCATCCCCCACTGAAGATGGATCATCCCACAGCAGCCCAGGACTGA
ATAAGCCATGCTCGCCCACCAGGCTTCTTTCTGACCCCTCATGTATCAAGCATCTCCAGGTG
ACCTACTGTCTGCCTATATTAAGGGTCCACACGGCGGCTGCTGCTTTTTTTTAAAGGGGAAA
GTAAGAAAAGAGAAGGAAATCCAAGCCAGTATATTTGTTTTATTTATTTTTTTTAAAGAAGAA
AAAAGTTCATCTTCACAAGGTGCTTCAGACTTGGTTTCTTAGCTAGAAACCAGAAGACTACG
GGAGGGAATATAAGGCAGAGAACTATGAGTCTTATTTTATTACTGTTTTTCACTACCTACTC
CCACAATGGACAATCAATTGAGGCAACCTACAAGAAAACATTTACAACCAGATGGTTACAAA
TAAAGTAGAAGGGGAAGATCAGAAAACCTAAGAAATGATCATAGCTCCTGGTTACTGTGGACT
TGATGGATTTGAAGTACCTAGTTCAGAACTCCCTAGTCACCATCTCCAAGCCTGTCAACATC
ACTGCATATTGGAGGAGATGACTGTGGTAGGACCCAAGGAAGAGATGTGTGCCTGAATAGTC
GTCACCATATCTCCAAGCTTCCTGGCAACCAGTGGGAAAAGAAACATGCGAGGCTGTAGGAA
GAGGGAAGCTCTTCCTTGGCACCTAGAGGAATTAGCCATTCTCTTCCTTATGCAAAGATTGA
GGAATGCAACAATATAAAGAAGAGAAGTCCCCAGATGGTAGAGAGCAGTCATATCTTACCCC
TAGATGTTTCATCCCAGCAGAAGAAAGAAGGTGTTGGGGTAGGATTCTTCAGAGGTTAGC
CTGGTACTTTCTCATCAGACACTAGCTTGAAGTAAGAGGAGAATTATGCTTTTCTTTGCTTT
TTCTACAAACCCTTAAAAATCACTTGTTTTAAAAAGAAAGTAAAGCCCTTTTCATTCAAAA
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

75/75

FIGURE 74

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA116510

><subunit 1 of 1, 494 aa, 1 stop

><MW: 54646, pI: 7.27, NX(S/T): 6

MARAAPLLAALTALLAAAAAGGDAPPGKIAVVGAGIGGSAVAHFLQQHFGPRVQIDVYEKGT
VGGRLATISVNKQHYESGAASFHSLSLHMQDFVKLLGLRHRREVVGRSAIFGGEHFMLEETD
WYLLNLFRLWWHYGISFLRLQMWWEEVMEKFMRIYKYQAHGYAFSGVEELLYSLGESTFVNM
TQHSVAESLLQVGVVTQRFIDDVVSAVLRASYGQSAAMPAFAGAMSLAGAQGSLWSVEGGNKL
VCSGLLKLTKANVIHATVTSVTLHSTEGKALYQVAYENEVGNSSDFYDIVVIATPLHLDNSS
SNLTFAGFHPPIDDVQGSFQPTVVSLVHGYNSSYFGFPDPKLFPPFANILTTDFPSFFCTLD
NICPVNISASFRRKQPQEAAVWRVQSPKPLFRTQLKTLFRSYYSVQTAEWQAHPLYGSRPTL
PRFALHDQLFYLNLEWAASSVEVMAVAANKVALLAYNRWYQDLDKIDQKDLMHKVKTEL

Important features of the protein:

Signal peptide:

amino acids 1-19

N-glycosylation sites.

amino acids 185-189, 290-294, 308-312, 312-316, 342-346, 378-382

N-myristoylation sites.

amino acids 33-39, 35-41, 38-44, 61-67, 64-70, 218-224, 234-240,
237-243, 429-435